BF THE UNIVERSITY OF MICHIGAN
OCT 2 1961
PERIODICAL READING ROOM

VOLUME 6 • NUMBER 3 • JULY 1961

CONTENTS

177 A Descriptive Approach to the Dynamics of Choice Behavior *by W. K. Estes*

185 Cross-Cultural Use of the Semantic Differential *by Howard Maclay and Edward E. Ware*

191 Testing Case-Descriptive Models *by William H. Starbuck*

200 Medical and Statistical-Psychological Models for Research in Schizophrenia *by Wilson H. Guertin*

CRITIQUE AND COMMENT

205 Personal Identity and Political Ideology *by Lucian W. Pye*

222 Categories, Concepts, and Reasoning in the Study of International Relations *by Richard N. Rosecrance*

232 Explorations in the Realm of Organization Theory. IV: The Simulation of Social Organization *by Richard L. Meier*

249 Book Reviews

COMPUTERS IN BEHAVIORAL SCIENCE

Symposium on The Use of Computers for Psychological Research

252 Introduction: Recent Areas of Application *by John E. Coulson*

252 The Computer as a Pattern Generator for Perceptual Research *by Benjamin W. White*

259 A Computer-Controlled Teaching Machine *by Harry F. Silberman*

261 Computer Simulation of Interaction in Small Groups *by A. Paul Hare*

265 Abstract of Empirical Inputs in a Voting Model *by W. N. McPhee*

266 Comments on the Use of Computers in Psychological Research *by Nancy S. Anderson*

271 Abstracts of Current Literature with Interdisciplinary Implications

275 About the Authors

BEHAVIORAL SCIENCE

A quarterly journal appearing in January, April, July, and October
Official Publication of the Mental Health Research Institute, The University of Michigan. Published at Mt. Royal & Guilford Aves., Baltimore 2, Maryland and the Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan

Behavioral Science contains articles on general theories of behavior and on empirical research specifically oriented toward such theories. An interdisciplinary approach to problems of behavior is stressed. Although the scope of the journal will include all aspects of behavior which can be subsumed under broadly general interdisciplinary theory, in the field of application special emphasis is placed on contributions relating to research in mental health and disease.

EDITORIAL BOARD: *Franz Alexander*, psychoanalysis—*Alex Bavelas*, social psychology—*Karl W. Deutsch*, political science—*David Easton*, political science—*Ralph W. Gerard*, biology (neurophysiology)—*Marion J. Levy, Jr.*, sociology—*Donald G. Marquis*, psychology—*Jacob Marschak*, economics—*James G. Miller*, psychiatry and psychology—*George P. Murdock*, anthropology—*Theodore M. Newcomb*, social psychology—*Anatol Rapoport*, mathematical biology—*Ralph W. Tyler*, education—*Raymond W. Waggoner*, psychiatry

Manuscripts should be sent in care of Dr. James G. Miller, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan

COMPUTERS IN BEHAVIORAL SCIENCE EDITORIAL COMMITTEE: *Steven G. Vandenberg*, University of Louisville—*Bert F. Green*, Massachusetts Institute of Technology—*Bruce MacKenzie*, International Business Machines—*Charles Wrigley*, Michigan State University.

Manuscripts for the Computers in Behavioral Science department should be sent in care of Dr. Steven G. Vandenberg, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan.

Subscription price: $6.00 a year. Price per issue: $1.75
Send subscriptions, address changes, and all business communications to *Behavioral Science*, Mental Health Research Institute, The University of Michigan, Ann Arbor, Michigan. Make all checks payable to *Behavioral Science.*

Second-class postage paid at Ann Arbor, Michigan and at additional mailing offices.
Copyright, 1961, by the Mental Health Research Institute, The University of Michigan.

To one interested in how an individual chooses among alternative courses of action, the motley array of theories of choice behavior must seem bewildering. Questions of very different sorts may be raised about the act of choice, and these give rise to quite different kinds of investigations. Does one wish to know more about how choice ought to be made? Or does one's interest lie in explaining behavior in terms of models representing observed choices as the result of postulated mental or neurophysiological events? What about the prediction of choice, the relating of decision probabilities to behavioral correlates and to observed independent variables? Does learning play a central role in choice behavior? Presented here is an attempt to organize several relevant viewpoints in terms of the kinds of questions they try to answer and to point toward a path of simplification.

A DESCRIPTIVE APPROACH TO THE DYNAMICS OF CHOICE BEHAVIOR*

by W. K. Estes

Indiana University

For the philosopher, social scientist, or other "outsider" who would gain an acquaintance with theories of choice behavior, the number and variety of current approaches must be overwhelming. The situation is little better even for the psychologist, accustomed though he is to a multiplicity of theories competing for attention in relation to almost every problem. For one must be prepared not only to accommodate a diversity of viewpoints, but to muster a not inconsiderable technical background. In some instances, current theories of choice involve rather formidable mathematical apparatus; in some instances they are associated with extensive programs of experimental research. I would, if it were possible, wish to proceed in this paper, first to answer the natural question as to why such a patchwork of choice theories should have arisen and flourished, then to show how all will ultimately be combined into a harmonious and manageable whole. What I shall actually attempt is first to organize the array of contemporary theories in terms of the kinds of questions they try to answer, then to offer some reasons to believe that continued development of one type of theory may eventually lead to simplification by eliminating the *raison d'être* for some of the currently fashionable varieties of choice theory.

* Reprinted, with permission, from *Logic, methodology and philosophy of science*. Proceedings of the 1960 International Congress, edited by Ernest Nagel, Patrick Suppes, and Alfred Tarski. Stanford, Calif.: Stanford University Press, 1961.

ALTERNATIVE APPROACHES TO CHOICE BEHAVIOR

To provide an initial overview of the area we may assemble the various approaches in outline. In all instances the object of attention is the behavior of an individual in choosing among alternative courses of action. But questions of very different sorts may be raised about the act of choice, and these give rise to quite different kinds of investigation. Questions as to how choices ought to be made have led to systems of the normative type. Calls for explanations of choice behavior have led, initially at least, to models representing observed choices as the resultant of postulated mental or neurophysiological events. Interest in the prediction of choice has instigated development of psychological scales and other models relating choice probabilities to behavioral correlates and to

observed independent variables. Thus we have:

I. Normative theories
   A. Moral (ethics)
   B. Rational (theory of games and statistical decision)
II. Quasi-normative theories (minimization of regret, maximization of expected value)
III. Hypothetical models
   A. Cognitive and purposive (faculty psychology)
   B. Analogical (physical models, computer simulation)
   C. Neurophysiological
IV. Descriptive theories
   A. Static (scaling and linear programming models)
   B. Dynamic (single and multiple process learning theories).

Perhaps because man's need to replace the apparent capriciousness of human actions with some kind of order so long preceded the availability of techniques for the scientific study of behavior, it has been a conspicuous characteristic of disciplines dealing with human choices that the elaboration of theory has far outrun the front line of empirical investigation. The precedent set by the great deductive systems of ethics has been more than upheld in the modern systems of game theory and statistical decision theory. These latter are presented formally as purely normative theories, designed to yield rules for optimal decision making, given specific assumptions as to the state of nature and criteria for determining optimality (see, e.g., Luce and Raiffa, 1958).

The popularity of normative theories is doubtless due in part to the notion of free will firmly ingrained in our everyday language. If an individual's choices were free of any external influence and if he knew the rules for best making decisions so as to achieve his purposes, then no further theory would be required to account for his choices. But only a few theorists well versed in counterfactual conditionals are able to entertain this possibility seriously. Aside from the difficulty that the premises are sharply contrary to both biological and psychoanalytic psychology, a theory limited to normative assumptions could not even attempt an interpretation of the obvious fact that a great proportion of human choices are far from optimal as judged by any recognized standard of values.

Nevertheless, people, theorists included, like to maintain the belief that they are basically rational even though their actions may not always appear so. When an individual's choices are not optimal, it may be that he has been conforming faithfully to the decision rules of a normative theory but has gone astray because of imprecision in his assignment of probabilities to uncertain events or of values to possible outcomes of his choices. Following this line of thought, social scientists have found it easy to generate plausible choice models almost ad infinitum by means of minor emendations of well-elaborated normative theories (Edwards, 1954; Thrall, Coombs, & Davis, 1954). One popular branch of this family of models stems from the assumption that individuals make choices so as to maximize the expected value of their outcomes in the long run (Edwards, 1955). Other variants are based on principles of minimum regret (Edwards, 1956), minimax regret (Savage, 1951), and so on. In each case the individual is conceived to function as a modernized version of the rational man of classical economics. His choices simply reflect his persistent efforts to maximize or minimize the expected value of some quantity or combination of quantities. The necessary empirical slack is provided by defining expected value in terms of the individual's subjective scales of utility or probability. The question as to how the individual makes use of past experience in achieving his goals of maximizing or minimizing has not been dealt with by theorists in the normative tradition until very recently. In an unpublished study by Radlow and Siegel (1960), an attempt is made to answer the question by assuming that subjective probability develops in accordance with the axioms of statistical learning theory.

Even more sharply than the normative approach, if this be possible, the more clearly hypothetical models illustrate the extent to

which theories of choice are determined by the background of the theorist rather than by factual considerations. With the growing popularity of computer simulation of mental processes, it is only to be expected that the older conceptualizations of choice in terms of mental faculties will be supplanted by computer models. A development of this sort is Toda's "Brain-computer" model (Toda, 1961) which combines computer simulation with normative considerations. Although the notion is understandably foreign to social scientists, some biologically oriented theorists (e.g., Deutsch, 1960) maintain that the only legitimate function of a behavioral theory is to yield information about underlying structures. Reflecting this view, Deutsch (1961) proposes to account for observed choice behavior in terms of hypothetical neurophysiological processes.

Although the custom of treating choice behavior has been first to select a preferred type of theory and then to seek some kind of empirical support for it, there would seem to be no logical reason why one could not proceed as with more neutral phenomena, beginning with experimental analysis and introducing as needed assumptions and concepts of whatever type that prove suitable for describing and predicting empirical data. In fact, an important example of the descriptive approach appeared as early as 1945 in Thurstone's application of the Law of Comparative Judgment to the prediction of choice. Owing to their strictly static character, Thurstone's models could not by themselves attempt a generally satisfying interpretation of choice behavior, but they have provided the basis for the scales of subjective probability and utility that provide some link between the normative theories and observed phenomena. Several currently active lines of research are like Thurstone's in that the models generated serve efficiently the function of data-reduction for certain types of choice experiments but do not attempt to relate choice behavior to antecedent conditions (e.g., Davidson, Suppes, & Siegel, 1957; Luce, 1959).

A descriptive theory as closely geared to data as the scaling models but capable of representing relationships between changes in choice tendencies and determining variables would fill the one conspicuous gap in the array of choice theories. Current trends within the area of choice and decision making do not, however, appear congenial to any such development. Learning theories are by their nature equipped to deal with changing response probabilities, but they have been developed primarily in conjunction with simple conditioning situations; and learning axioms stated in terms of stimulus and response probabilities cannot lead directly to statements about optimal strategies, utility scales, or other like concepts prominent in the literature of choice and decision.

Against the weight of tradition are two recent developments pointing to possible links between learning and choice theory. One of these is Luce's demonstration that the existence of such psychological scales as that of utility can be derived from axioms stated solely in terms of response probabilities (Luce, 1959). The other is the application of statistical learning models to situations in which the individual is confronted with repeated opportunities to choose among a given set of alternatives, learning being manifested by systematic changes in choice probabilities over trials as a function of values or probabilities of the trial outcomes (Bush & Estes, 1959; Estes, 1957; Grant, Hake, & Hornseth, 1951; Suppes & Atkinson, 1960). In the remainder of this paper, we shall explore the potentialities of this kind of development by summarizing the assumptions of one of these learning models and illustrating some results of its application to binary choice experiments.

ELEMENTS OF A LEARNING THEORY FOR BINARY CHOICES

The particular theory to be discussed represents an interpretation of the simplest of the statistical learning models, the "pattern model" (Bush & Estes, 1959, Ch. 1). In this theory, a series of choice trials is conceptualized in terms of the stimulus situation (pattern) which can prevail at the beginning of a trial, the response alternatives among which the subject must choose, and the trial outcomes, whose reinforcing

effects serve to establish learned associations between stimulus patterns and responses.

For our first illustrative interpretation, we shall consider a situation which involves uncertainty with respect to trial outcomes, but no differences in utility for different outcomes, viz., the familiar experiment on simple predictive behavior or verbal conditioning (Estes & Straughan, 1954; Grant et al., 1951; Neimark, 1956). The subject's task is simply to choose between two response alternatives, e.g., to operate one of two response keys or to say one of two designated words at the beginning of each trial, attempting to be "correct" as often as possible. The trial terminates with a signal indicating to the subject which response was "correct" according to the experimenter's protocol. Typically, the sequence of signals is programmed by the experimenter so that one response has some fixed probability π and the other probability $1 - \pi$ of being correct on any trial.

In conceptualizing this experimental situation, we assume that the experimental conditions determine a set of N stimulus patterns any one of which may be present at the beginning of any trial, and all of which are equally likely to occur. A stimulus pattern includes any instruction or other signal given by the experimenter to indicate the initiation of a trial, and may include also whatever aspects of the preceding sequence of trial outcomes the subject has been able to verbalize and recall. Designating the response alternatives by A_1 and A_2, trial outcomes may be categorized into two classes, those that reinforce A_1 and those that reinforce A_2. The principal assumptions concerning the learning process are as follows.

1. Each stimulus pattern is at any time associated with ("connected to") exactly one of the responses A_1 or A_2.

2. When a trial begins with a stimulus pattern connected to A_i, A_i is chosen by the subject with probability one.

3. When a trial begins with a stimulus pattern connected to A_i and the outcome is one which reinforces A_i (indicates A_i was correct), the connection of the pattern to A_i remains unchanged.

4. When a trial begins with a stimulus pattern connected to A_i and the outcome is one which reinforces A_j ($j \neq i$), then with probability c the pattern becomes connected to A_j, while with probability $1 - c$ its connection to A_i remains unchanged.

PROBABILITY MATCHING

It is an obvious theorem that if k of the N patterns are associated with A_i, the probability of A_i is equal to k/N. Not quite so obvious, but still quite easily derived (Bush & Estes, 1959, Ch. 1; Estes, 1957), is the "matching theorem," which states that over a series of trials in which the probability of outcomes that reinforce A_i is constant and equal to π, the probability of an A_i choice on the part of the subject approaches π asymptotically. In other words, probability of choosing a given alternative tends to match its probability of reinforcement. There is, to be sure, an element of ambiguity in interpreting the notion of an asymptote with respect to any aspect of human behavior, for even in the best controlled experiment, the conditions influencing the subject's behavior cannot be held constant indefinitely. Nevertheless, the matching theorem has received rather striking confirmation in a number of experiments.

To give a concrete indication of the degree of confirmation, I have pooled the results reported for 18 groups of human subjects run in six experiments by different groups of experimenters during the period 1939–1956 (Estes & Straughan, 1954; Goodnow & Postman, 1955; Grant et al., 1951; Humphreys, 1939; Jarvik, 1951; Neimark, 1956). The π values for the more frequently reinforced response ranged from .50 to 1.00 with a mean of .76. The mean asymptotic probability of the more frequently reinforced response, estimated by the arithmetic mean of the response proportions for the terminal trial block in each experiment, was .76. The standard deviation of the distribution of differences between π values and mean response proportions was less than .02.

This phenomenon of "probability matching" has excited considerable general interest, partly, no doubt, simply because parameter-free[1] quantitative predictions

[1] By "parameter-free," we refer to predictions which do not depend on the evaluation of free parameters from data, but only on experimenter determined parameters, as π in the present instance.

have rarely been attempted, let alone confirmed, in psychology. Another source of interest has been the disparity between the matching result and the common preconception (evidently based partly on intuition and partly on injudicious extrapolation of game theory) that behavior in the two-choice situation should always conform to a normative principle of maximizing the expected frequency of correct responses. Unfortunately, the question of "matching vs. maximizing" cannot be answered simply by the examination of asymptotic data from choice experiments. Concerning the conclusion of the experiments cited in the preceding paragraph, the criticism was immediately offered that the experiments, typically involving learning series of 100 to 200 trials, might have been terminated too soon and that the observed matching of response and reinforcement frequencies might depend on a fortuitious selection of series duration. In several more recent studies (Anderson & Whalen, 1960; Cotton & Rechtschaffen, 1958; Gardner, 1957; LaBerge, 1959; Neimark & Shuford, 1959), learning series have been more than doubled in length (running from 300 to 450 trials) with no appreciable departures from the previous pattern of results. If one claimed that the asymptote of two-choice learning had now been evaluated, the same criticism could be raised again. It would, however, be hard to dispute the conclusion that over at least the initial stages of simple two-choice learning, probabilities of response tend to approach their probabilities of reinforcement.

DESCRIPTION OF SEQUENTIAL DATA

Current research in this area is not by any means preoccupied with the matter of asymptotic probability matching. Emphasis is, rather, shifting to the more ambitious task of accounting for detailed quantitative properties of behavior throughout the course of learning. If the assumptions of the statistical model are sound, it should be possible to predict probabilities of a given choice following any specified sequence of preceding responses and reinforcing events at any stage of learning. One illustration of the level of accomplishment now attainable is given in Table 1. The data come from a study of simple two-choice learning reported by Friedman, Burke, Cole, Estes, and Millward (1960), in which a group of 80 college student subjects, after having extensive preliminary training with a variety of different reinforcement probabilities (π values), was run through a series of 288 trials with $\pi = .80$. An entry $p_{ij}(k)$ represents the probability of response k on a given trial following the combination of response i and reinforcement j on the preceding trial averaged over trials 1–12 (initial block) or 241–288 (final block). The theoretical values (expressions for which are given in Bush and Estes [1959, Ch. 1]) are those predicted by the pattern model described above after the parameters of the model have been estimated from the data by a modified least squares procedure. Considering that the number of parameters is small relative to the number of degrees of freedom in the data, and that the value of c/N is the same for both blocks, a fairly stringent test of this aspect of the model is provided.

TABLE 1

OBSERVED AND THEORETICAL PROPORTIONS OF MORE FREQUENTLY REINFORCED RESPONSE IN TWO-CHOICE LEARNING EXPERIMENT WITH .80:.20 REINFORCEMENT PROBABILITIES

	Initial Block		Final Block	
	Obs.	Th.	Obs.	Th.
$p_{11}(1)$	.38	.38	.59	.58
$p_{21}(1)$	.17	.18	.10	.08
$p_{12}(1)$	.06	.07	.13	.12
$p_{22}(1)$	.04	.03	.02	.02
$p_{11}(2)$	.13	.13	.07	.06
$p_{21}(2)$	.12	.12	.04	.07
$p_{12}(2)$	.06	.05	.04	.04
$p_{22}(2)$	.04	.04	.02	.02

EXTENSION OF THE PATTERN MODEL TO PAIRED COMPARISON LEARNING

Since the assumptions of the statistical model include no direct mention of values or costs, it is not particularly surprising to find that when differential monetary incentives and risks are introduced into the two-choice situation, choice probabilities deviate from the theoretically predicted values. In studies by Siegel and Goldstein (1959) and by Suppes and Atkinson (1960), among

others, monetary rewards have been given for correct choices and fines for incorrect ones. The result, relative to asymptotic behavior, has been a tendency for terminal choice proportions to overshoot matching values, the amount of the disparity being directly related to the magnitudes of the incentives or risks. One interpretation of this finding is that the dynamics of the learning process are basically different in the situation with incentives and risks. A second possibility is that there are two kinds of learning going on in this situation: (1) changes in association between responses and stimulus patterns, produced by the individual trial outcomes in accordance with the assumptions of the pattern model; and (2) changes in habits or hypotheses controlling behavior over longer sequences of trials, arising from the individual's response to the larger swings in his cash balance. If the latter interpretation is correct then we might expect to find, somewhat paradoxically, that the simple model will again provide correct predictions if we go to a still more complex incentive-risk situation in which the subjects cannot readily identify and verbalize the relationships between their courses of action and mean values of outcomes over series of trials.

A timely opportunity to test this last suggestion is provided by a "paired comparison" learning experiment conducted in Professor Patrick Suppes' laboratory at Stanford University (for a brief description see Suppes and Atkinson [1960]). In this experiment the subject was offered a choice from a subset of three objects on each trial. Designating the three alternatives as A, B, and C, the combinations AB, AC, BC, and ABC were each presented on 100 trials of the 400 series in random sequence. The experimenter's program, unknown to the subject, assigned A, B, and C, respectively, probabilities of .67, .40, and .20 that a choice of the given alternative would be "correct" on any trial on which it was offered. For one group of 48 subjects each correct choice was rewarded with one cent and each incorrect choice cost the subject one cent; for a second group of 48 subjects, each correct choice was rewarded with five cents and each incorrect choice cost the subject five cents.

TABLE 2

OBSERVED AND PREDICTED TERMINAL CHOICE PROPORTIONS FOR PAIRED-COMPARISON LEARNING EXPERIMENT

	Observed		Predicted
	1¢ Gr.	5¢ Gr.	
$p_{AB}(A)$	.65	.70	.64
$p_{BC}(B)$	.56	.56	.57
$p_{AC}(A)$	.70	.71	.71

To apply the pattern model to this situation, we assume that each of the four presentation sets, AB, AC, etc., constitutes a stimulus pattern and that associations of choice responses with any one pattern are learned, over the sequence of trials on which that pattern is present, in accord with the model. As in the simpler case considered above, the model predicts asymptotic probability matching; but there is a complication in that the matching value is now under the joint control of the experimenter and the subject. Consider, for example, the pattern AB. Letting π_a and π_b represent the probabilities that choices of A and B, respectively, will be correct on any trial, one can show by a simple probabilistic argument (see, e.g., Bush and Estes, 1959, Ch. 1; Suppes and Atkinson, 1960; Thrall et al., 1954, Ch. 8 & 9), that asymptotically the probability of a choice of A will match its probability of reinforcement[2] only if both are equal to $(1 - \pi_b)/(2 - \pi_a - \pi_b)$ (which, of course, reduces to π in the special case, previously considered, where $\pi_a = 1 - \pi_b = \pi$). In Table 2 asymptotic choice probabilities computed from this expression are compared with observed proportions over trials 301–400 for the one cent and five cent groups.[3] Although a full evaluation of the predictive power of the model for this situation must wait upon more detailed analyses, including sequential statistics of the data (currently in progress), the results

[2] In this situation, a given choice response is said to be reinforced if it occurs and is correct or if the alternative response occurs and is incorrect.

[3] I am indebted to Professor Suppes for making available these unpublished data.

for terminal choice proportions in the binary subsets appear quite promising.

CONCLUDING REMARKS

In view of recent developments of the sort sketched in the preceding section, it can scarcely be doubted that encroachments of learning theory into the area of choice and decision will continue, probably on a steadily expanding scale. The potentialities and limitations of this approach can scarcely be gauged at the present stage. The denouement which I, personally, find most congenial to contemplate is an ultimate simplification of the currently motley array of overlapping theoretical efforts into two main kinds: strictly normative theory, essentially a branch of applied mathematics; and a system of descriptive models based on concepts and laws of learning and behavior theory—the latter subsuming the function of the current quasi-normative, faculty, and analogical models.

At best, progress toward such an outcome would necessarily be slow, for in this area theories which must make their way solely on their ability to describe and predict empirical data face resistance in depth from entrenched conceptions of man as a rational decision maker, as well as from some rather anachronistic views of theory which are prevalent among social and behavioral scientists. Typical of these last is the view expressed by Deutsch (1960; 1961) which would relegate to second-class citizenship theories which simply describe and predict data. While sharing Deutsch's admiration for theories which generate inferences about underlying structures, one can doubt the wisdom of our straining so hard for interpretation in terms of structure that we fail to discover the "laws of motion." If a functional, descriptive theory of choice proves dependable in helping us to analyze and anticipate the outcomes of new situations, we will probably find that it simultaneously satisfies our demand for an intuitive understanding of choice phenomena.

REFERENCES

Anderson, N. H., & Whalen, R. E. Likelihood judgments and sequential effects in a two-choice probability learning situation. *J. exp. Psychol.*, 1960, 60, 111–120.

Bush, R. R., & Estes, W. K. *Studies in mathematical learning theory.* Stanford, Calif.: Stanford Univ. Press, 1959.

Cotton, J. W., & Rechtschaffen, A. Replication report: Two and three choice verbal-conditioning phenomena. *J. exp. Psychol.*, 1958, 56, 96.

Davidson, D., Suppes, P., & Siegel, S. *Decision making: An experimental approach.* Stanford, Calif.: Stanford Univ. Press, 1957.

Deutsch, J. A. Learning-behavior-information. *Contemp. Psychol.*, 1960, 5, 147–150.

Deutsch, J. A. The implications of an experiment for theories of choice. *Proceedings of the 1960 International Congress for Logic, Methodology and Philosophy of Science.* Stanford, Calif.: Stanford Univ. Press, 1961.

Edwards, W. The theory of decision making. *Psychol. Bull.*, 1954, 51, 380–417.

Edwards, W. The prediction of decisions among bets. *J. exp. Psychol.*, 1955, 50, 201–214.

Edwards, W. Reward probability, amount, and information as determiners of sequential two-alternative decisions. *J. exp. Psychol.*, 1956, 52, 177–188.

Estes, W. K. Of models and men. *Amer. Psychologist*, 1957, 12, 609–617.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical learning theory. *J. exp. Psychol.*, 1954, 47, 225–234.

Friedman, M. P., Burke, C. J., Cole, M., Estes, W. K., & Millward, R. B. Extended training in a non-contingent two choice situation with shifting reinforcement probabilities. Paper given at 1960 meetings of the Psychonomic Society.

Gardner, R. A. Probability learning with two and three choices. *Amer. J. Psychol.*, 1957, 70, 174–185.

Goodnow, J. J., & Postman, L. Probability learning in a problem solving situation. *J. exp. Psychol.*, 1955, 49, 16–22.

Grant, D. A., Hake, H. W., & Hornseth, J. P. Acquisition and extinction of a verbal conditioned response with varying percentages of reinforcement. *J. exp. Psychol.*, 1951, 42, 1–5.

Humphreys, L. G. Acquisition and extinction of verbal expectations in a situation analogous to conditioning. *J. exp. Psychol.*, 1939, 25, 294–301.

Jarvik, M. E. Probability learning and a negative recency effect in the serial anticipation of alternative symbols. *J. exp. Psychol.*, 1951, 41, 291–297.

LaBerge, D. Effect of preliminary trials on rate of conditioning in a simple prediction situation. *J. exp. Psychol.*, 1959, 57, 20–24.

Luce, R. D. *Individual choice behavior.* New York: Wiley, 1959.

Luce, R. D., & Raiffa, H. *Games and decisions.* New York: Wiley, 1958.

Neimark, E. D. Effects of type of non-reinforcement and number of alternative responses in two verbal conditioning situations. *J. exp. Psychol.*, 1956, 52, 209–220.

Neimark, E. D., & Shuford, E. H. Comparison of predictions and estimates in a probability learning situation. *J. exp. Psychol.*, 1959, 57, 294–298.

Radlow, R., & Siegel, S. Decision making and learning. Department of Psychology Research Bulletin No. 6 (mimeo). State College, Pa.: Penn State Univ., 1960.

Savage, L. J. The theory of statistical decisions. *J. Amer. Statistical Assoc.*, 1951, 46, 55–67.

Siegel, S., & Goldstein, D. A. Decision-making behavior in a two-choice uncertain-outcome situation. *J. exp. Psychol.*, 1959, 57, 37–42.

Suppes, P., & Atkinson, R. C. *Markov learning models for multiperson interactions.* Stanford, Calif.: Stanford Univ. Press, 1960.

Thrall, R. M., Coombs, C. H., & Davis, R. L. (Eds.) *Decision processes.* New York: Wiley, 1954.

Thurstone, L. L. The prediction of choice. *Psychometrika*, 1945, 10, 237–253.

Toda, M. "Brain-computer" model approach to the theory of choice. *Proceedings of the 1960 International Congress for Logic, Methodology and Philosophy of Science.* Stanford, Calif.: Stanford Univ. Press, 1961.

(Manuscript received January 18, 1961.)

"The more one contemplates our own time, the more it seems to resemble those earlier dynamic and critical eras of Western history so typically characterized by a keen sense of change and exploration, and also by an atmosphere of personal and political and social insecurity. No one observing the eagerness with which developments in the exploration of space, for example, are currently watched and commented upon and interpreted by all the peoples of the globe can forget the world of Leonardo, into which Columbus and Magellan and Vasco da Gama were born. Our time is marked by the same sense of new physical frontiers close at hand, the same ferment of new ideas, the same vast and rending shifts in the individual's notion of the universe and his place in it, and, most poignantly, of his very nature."

CARYL P. HASKINS, *The Report of the President of Carnegie Institution of Washington.* 1959–1960

Osgood's semantic differential is a way to discover where any word is located in a conceptual space whose dimensions are meanings rather than real-world distances. It seems likely that the space of meanings would differ from culture to culture just as it probably does somewhat from individual to individual. Here is an example of interdisciplinary research—the use of a proven psychological measuring instrument such as the semantic differential with data that have traditionally been the concern of anthropology. In addition to expanding the range of research techniques for another discipline, such cross-cultural studies are of value for a fuller understanding of the limits within which the instrument is effective; they refine and improve research methods.

CROSS-CULTURAL USE OF THE SEMANTIC DIFFERENTIAL

*by Howard Maclay and Edward E. Ware**

Institute of Communications Research, University of Illinois

A CENTRAL problem in research is the extent to which the methods and conclusions of one discipline can be applied to the data of another. We are concerned here with methods of gathering and interpreting data rather than with the transfer of general theories from one area to another. We propose to take a measuring instrument developed by experimental psychologists, the semantic differential (Osgood, Suci, & Tannenbaum, 1957), and apply it to data that have traditionally been the primary concern of anthropologists. A double value may be derived from such an approach. Psychological instruments have most often been developed and validated in the context of Western culture. An important question involves the extent to which conclusions based on these instruments can be extended to human beings generally. This can only be studied by applying the instruments to a subject population consisting of persons with non-Western cultural backgrounds. An inevitable result of this approach is a fuller understanding of the limits within which the instrument is effective. The extensive use of the Rorschach test by anthropo'ogists (summarized by Henry and Spiro, 1953) is an illustration of how a psychological technique can be refined and improved by cross-cultural application.

Anthropologists will also benefit if a technique developed elsewhere proves to have cross-cultural validity and produces results which bear on problems of culture structure and process. Traditional anthropological field methods, which commonly involve some sort of participant observation, are extremely laborious and time consuming. If one wishes to describe a functioning culture, there is clearly no substitute for an extended period of systematic observation. It may be, however, that some part of the data needed for an adequate description can be gathered more effectively by other methods.

Our focus here is on the possible utility of a particular measuring device for anthropologists. This implies first that its cross-cultural validity must be demonstrated and further that its usefulness as a specific anthropological tool be shown. A fully convincing demonstration of validity is not possible in a short paper. This can only follow from extensive use in the field. Our aim here is to present data which will contribute to a final decision as to the

* This research was fully supported by the Southwest Project in Comparative Psycholinguistics sponsored by the Social Science Research Council and directed by John B. Carroll. Research teams responsible for the data upon which the present analysis is based included J. Casagrande, S. Ervin, A. Horowitz, and H. Landar for Navaho; J. Bastian, S. Bastian, E. Lenneberg, and S. Newman for Zuni; J. Carroll, L. Elliott, and H. Maclay for Hopi. We are indebted to Harry Basehart, Joseph B. Casagrande, and Charles E. Osgood for helpful comments on a first draft of this paper.

anthropological worth of the instrument in question along with some observations based on previous use of the semantic differential in a number of field situations. The estimation of validity requires that evaluation be based on criteria which are independent of the instrument being judged. We will regard an instrument as having cross-cultural validity to the extent that it reflects cultural differences known to exist as a result of previous ethnographic study. We first discuss the instrument and then the cultural criteria against which it is to be judged. An application to subjects from three American Indian cultures is described followed by a concluding section in which an estimate of the anthropological value of the instrument is presented.

THE PSYCHOLOGICAL INSTRUMENT

The measuring device under investigation is a form of the semantic differential developed at the University of Illinois by C. E. Osgood and associates and fully described by Osgood, Suci, and Tannenbaum (1957). The semantic differential is a technique for measuring the meaning of linguistic forms. Concepts, such as *mother* or *President Eisenhower*, are rated on a series of descriptive scales. Each seven point scale is defined by a pair of opposite adjectives (i.e., good-bad, hot-cold, active-passive, etc.). If we are interested in the meaning of the English word *man*, we might present native English speakers with the following form:

MAN

good___:___:___:___:___:___:___bad
strong___:___:___:___:___:___:___weak
fast___:___:___:___:___:___:___slow

It should be noted that this is a method of measurement and not a specific test. Particular forms of the semantic differential may vary in the concepts and scales used in terms of the research purposes of the investigator.

Subjects indicate the direction and intensity of the association between the concept and the adjectives defining each scale. The meaning of a concept is given by the positions it occupies on a set of scales. Factor analyses of semantic differential results have led to the establishment of three main factors which are regarded as the primary dimensions that organize semantic judgments. These are named *evaluative* (good-bad), *potency* (strong-weak), and *activity* (fast-slow). Any concept may be located in a semantic space defined by these factors.

A final assessment of the cross-cultural validity and utility of the semantic differential must rest on the answers to two questions: (1) Is there some set of basic semantic dimensions that is universal? (2) If this is the case, can the instrument organize and classify cultures in a scientifically interesting way? The dependence between these problems is obvious. We can compare cultures directly only if the dimensions of comparison are universal. On the other hand, a similarity in factor structure across cultures does not imply that the meanings assigned to individual words are the same but only that the scales are used in an equivalent way. While the first problem has not been finally settled, a number of studies have found factor structures the same or quite similar to English for Japanese and Korean (Kumata & Schramm, 1956; Kumata, 1957), Greek (Triandis & Osgood, 1958), and Hopi, Zuni, Navaho, and Spanish-American (Suci, 1960).[1] Work in this area has progressed to the point where an exploratory study addressed to the second question seems profitable. We will assume the universality of the semantic space and focus our attention upon the specific meanings assigned to concepts by our subjects.

THE CULTURAL CRITERIA

We now consider the cultural distinctions which will serve as criteria in evaluating this instrument. These arise from the observational methods of cultural anthropology and are thus methodologically independent of any results which the semantic differential may produce. An effort has been made to select cultural distinctions that are compelling and obvious to the point where all anthropologists familiar with the cultures involved would agree that they do indeed exist. We then ask whether the semantic differential is able to reflect these

[1] C. E. Osgood is currently directing a project whose aim is to specify more precisely a set of universal dimensions.

known distinctions. We will regard as clearly distinct any two cultures whose ethnographic descriptions differ and whose members speak mutually unintelligible languages. This means that a significant proportion of translation equivalent forms should be located differently in the semantic space for the two cultures. More specifically, most of the concepts used should have different locations on the scales of the semantic differential. It is assumed that the scales are being used in the same way and that direct comparisons of the concepts can be made. The three American Indian groups studied are Hopi, Zuni, and Navaho. They are known to be distinct cultures and we ask if the instrument can distinguish them. If the distinction is made, it can then be asked if the instrument orders the three cultures in an acceptable way. Hopi and Zuni are usually classified as Western Pueblo (Eggan 1950) and it was therefore predicted in advance of the analysis that they should fall together as opposed to Navaho. It was further predicted that Zuni should be closer to Navaho than Hopi on the less secure grounds of more formal and more amicable Zuni-Navaho contact than Hopi-Navaho as described by Vogt (1951).

SUBJECTS

The Hopi group consisted of 28 subjects, the Navaho of 29 subjects, and the Zuni of 26 subjects. The groups ranged in age from the middle teens to over 60 and included both males and females.

MATERIALS

The following form of the semantic differential was used:

Concepts	*Scales*
1. Coyote	1. light-heavy
2. Mexican	2. hard-soft
3. Corn	3. ugly-pretty
4. Male	4. cold-hot
5. Horse	5. dirty-clean
6. Rain	6. wet-dry
7. Female	7. industrious-lazy
	8. sad-happy
	9. poor-rich
	10. sweet-sour
	11. short-long
	12. fast-slow
	13. strong-weak
	14. crooked-straight
	15. good-bad

PROCEDURE

The semantic differential was translated into the three languages by linguists working with informants (c.f. Ervin [1956] for a discussion of translation procedures). The typical procedure using a seven-point scale described above was modified somewhat by the requirements of the field situation. Suci (1960) has described the exact procedures and the following is a paraphrase of his discussion.

The subjects were nonliterate in their own language and had to be given the material verbally and individually. Preliminary investigation had indicated that the standard rating procedure was not understood by subjects and might have led to ambiguous results. Consequently, a six-inch line with the ends and middle marked was used for each scale. The order of the scales and of the concepts, and the polarity of the modifiers (e.g., *good-bad* vs. *bad-good*) was determined by randomization. The differentials were scored by measuring the distance, in one-tenth inches, from the left side of the scale to the check mark locating the concept on the scale. Thus, a concept could vary from zero to sixty on any single scale.

It appeared that almost all subjects readily grasped the nature of the task and performed it satisfactorily.

RESULTS

To simplify interpretations and limit the number of variables in any particular comparison, Hotelling's T^2 (Hodges, 1955) was obtained from the 15 scales for each concept. The three cultures were compared by pairs. Hotelling's T^2 is the multivariate generalization of Student's t test. The t test is used to determine the statistical significance of the difference between the means of two groups of subjects on a single variable. T^2 is a statistical test of the difference between two groups on a set of several variables. The results of these comparisons are found in Table 1. Of the 21 comparisons, 17 show significant differences beyond the .05 level. We conclude on the basis of these results that the instrument can discriminate among cultures.

TABLE 1
F VALUES DERIVED FROM HOTELLING'S T^2 TEST OF THE DIFFERENCES BETWEEN PAIRS OF CULTURES FOR EACH CONCEPT

Concept	Hopi-Zuni	Zuni-Navaho	Hopi-Navaho
1. Coyote	1.28	3.79**	3.74**
2. Mexican	1.76	1.54	1.80
3. Corn Ear	2.30*	4.96**	4.54**
4. Male	2.89**	2.64**	6.05**
5. Horse	4.50**	2.12*	4.54**
6. Rain	4.50**	4.83**	2.56**
7. Female	3.46**	2.24*	4.54**

* Difference is statistically significant beyond the .05 level.

** Difference is statistically significant beyond the .01 level.

Individual comparisons were made between pairs of cultures for each concept-scale combination using the *t* test. These results, presented in Table 2, give a detailed picture of the culture-pair comparisons. Here, 94 of the 315 comparisons show a significant difference. Due to the large number of statistical tests based on a single sample, interpretation of the results should be considered as hypothesis gathering rather than hypothesis testing. Any specific findings require some form of cross-validation.

It can be observed that individual concepts and scales vary greatly in their ability to discriminate. The concept *coyote*, for example, seems to have a similar meaning for Zuni and Hopi and a different and distinctive meaning for Navaho. We have not explored in detail the cultural correlates of these results but students of the three cultures may find it profitable to attempt this sort of study. With the aim of facilitating such an investigation we present the actual means of all concepts on all scales for each culture in Table 3.

No statistical test of the ordering of cultural differences was available for this data. However, an intuitive measure of cultural distance may be derived by adding the number of significant concept-scale differences under each culture comparison in Table 2. There are a total of 23 significant differences between Hopi and Zuni, 34 between Navaho and Zuni, and 37 between Hopi and Navaho. If these numbers are taken as measures of cultural distance we can conclude that the trend in the data is in the predicted direction: Hopi and Zuni tend to fall together as against Navaho; Zuni is

TABLE 2
RESULTS OF *t* TESTS* OF THE DIFFERENCES BETWEEN CULTURE-PAIRS FOR EACH CONCEPT-SCALE COMBINATION

Scales**	Concepts** 1	2	3	4	5	6	7
Hopi-Zuni Comparisons							
1.					2.14		
2.							
3.							
4.						−4.03	
5.			2.09			3.87	
6.							
7.							
8.							
9.		−2.87	−2.43	−3.50		2.27	−3.50
10.						−4.48	
11.		−2.96			2.09		−2.97
12.			2.21		−2.42		4.37
13.							2.40
14.					−4.24		−3.98
15.	2.66	3.39	3.51		2.16		
Zuni-Navaho Comparisons							
1.	3.47					−4.53	
2.							
3.			−3.69	−2.84		−4.52	
4.					2.74		4.25
5.	2.08					−4.83	3.32
6.			3.94				
7.		−2.68		−2.31			
8.		2.11					
9.	3.24	2.71	3.49	2.22			3.06
10.	−2.78	−2.88				2.02	
11.	2.85						
12.			−2.07	−3.64	−3.11	−2.82	−3.24
13.	−5.83	−2.42					−3.27
14.	4.08						
15.	−2.53	−3.57					
Hopi-Navaho Comparisons							
1.	3.74					−2.43	
2.		−2.38	−2.38	−2.92	−2.34		−2.29
3.	2.35		−2.86	−2.76		−2.90	
4.						−2.72	3.88
5.			2.19	2.15			3.17
6.			4.76				
7.				−2.57			−2.55
8.							
9.	3.33						
10.	−2.03					−2.04	
11.	2.31						
12.				−4.26	−6.21	−2.85	
13.	−5.62	−3.09		−2.66	−3.39		
14.	2.67			−2.95	−4.83	−2.26	−3.75
15.			3.71				−2.38

* Only those values significant beyond the .05 level are given. Values greater than 2.68 are significant beyond the .01 level.

** Scales and concepts are in the same order as they are given in the section on materials.

TABLE 3

SEMANTIC DIFFERENTIAL SCALE MEANS* FOR ZUNI (Z), HOPI (H), AND NAVAHO (N) CULTURES**

Scales	Concepts																					
	1 Coyote			2 Mexican			3 Corn			4 Male			5 Horse			6 Rain			7 Female			
	Z	H	N	Z	H	N	Z	H	N	Z	H	N	Z	H	N	Z	H	N	Z	H	N	
1. Light	28	29	17	37	33	36	26	24	21	39	40	39	39	48	45	22	32	43	37	32	36	Heavy
2. Hard	24	27	30	26	23	31	28	23	33	28	21	32	25	16	25	37	32	32	30	27	35	Soft
3. Ugly	21	25	16	37	34	31	36	39	47	36	37	44	37	39	41	31	37	49	38	38	42	Pretty
4. Cold	33	36	32	37	34	35	22	25	19	36	33	35	42	39	34	23	9	19	40	38	26	Hot
5. Dirty	26	24	18	35	31	31	39	46	39	37	40	34	34	38	34	31	46	48	44	43	33	Clean
6. Wet	31	28	34	26	26	29	33	36	19	30	24	26	27	27	24	13	7	10	28	24	23	Dry
7. Industrious	30	34	26	19	24	30	21	18	20	17	16	24	18	17	20	21	19	16	18	15	22	Lazy
8. Sad	27	30	23	40	36	33	43	38	39	42	40	39	41	42	39	39	43	45	43	40	39	Happy
9. Poor	21	18	9	39	29	29	42	34	28	45	34	36	36	31	32	32	41	41	42	32	32	Rich
10. Sweet	33	36	43	24	28	34	16	13	17	23	25	22	26	19	21	25	9	16	20	20	24	Sour
11. Short	23	20	13	37	27	32	29	28	26	33	35	34	30	38	35	34	34	38	33	23	28	Long
12. Fast	17	16	23	24	25	28	24	34	32	19	18	32	17	10	29	20	20	33	21	37	33	Slow
13. Strong	21	23	43	19	18	28	22	25	24	17	14	21	14	9	17	18	13	17	17	26	28	Weak
14. Crooked	32	27	18	37	33	33	40	39	39	37	30	39	40	25	40	39	32	42	42	29	40	Straight
15. Good	33	45	44	17	33	30	14	31	16	17	16	22	16	26	19	14	19	14	15	14	21	Bad

* Means were computed from a 60-point scale so that low values are toward left-hand scale term (e.g., light) and high toward right-hand term (e.g., heavy).

** *t* values for significant culture-pair differences are given in Table 3. Concepts and scales are numbered in this table to facilitate use of Tables 2 and 3 together.

closer to Navaho than is Hopi. The lack of any significance test makes this conclusion suggestive rather than final.

DISCUSSION

We have considered here only one of the several criteria that an instrument must meet before it is acceptable as an anthropological field tool. With respect to its ability to reflect gross cultural divisions, the status of the semantic differential is promising. Zuni, Hopi, and Navaho were distinguished at a statistically significant level and they were also grouped internally as predicted.

Some advantages of the instrument lie outside the analyses performed for this study. The task required by this technique is a straightforward one that can be easily understood and accomplished by most subjects regardless of wide variations in their age, intelligence, and degree of acculturation. Even if the tests are administered orally and individually an impressive amount of data can be collected in a relatively short time; Hopi subjects marked scales at the rate of about 250 per hour. The tests can be administered by moderately trained persons and need not occupy the time of the senior investigator. The techniques of recording and analyzing data are objective and standardized although interpretation of the results does, of course, rest on the judgment of the analyst.

An obvious application of such an instrument is to problems of covert culture. Most definitions of culture include nonobservable elements such as values, ideas, and beliefs. The semantic differential has the potential of serving both to gather data in these areas and to evaluate hypotheses derived from other data. The evaluative factor, for example, offers a way of ranking events, people, and ideas on a good-bad continuum.

Any research device may have either full or auxiliary status in the methodological repertoire of a discipline. Full status implies that it is an essential data gathering or analyzing technique whose results are to be weighted equally with those obtained by other methods. An auxiliary method has the function of supplementing more basic techniques. It may provide detail in areas whose general outlines are established on other grounds. In this case its results tend to be checked against other information and

discarded if they vary too widely. It may also act as an opening wedge for the introduction of more highly regarded procedures. As an example of this function, the semantic differential tends to provoke little anxiety and puts the field worker in friendly contact with many members of the community in a short time. If the local group is bilingual its members are very much aware of linguistic differences and likely to become quite interested in problems of meaning in different languages. That the semantic differential was administered to a great many subjects in areas where field work is often extremely difficult (Hopi and Zuni) is testimony to its field efficiency.

We conclude that while this instrument has unquestioned merit as a general scientific technique its full acceptance as an anthropological method must await further research. We believe it can now function as an auxiliary method and, further, that it shows promise of becoming a useful device in studies of covert culture.

REFERENCES

Eggan, F. *Social organization of the western Pueblos.* Chicago: Univ. of Chicago Press, 1950.

Ervin, S. Translation procedures. In J. B. Carroll and S. M. Ervin (Eds.), *Field manual of the Southwest Project in Comparative Psycholinguistics.* Cambridge: Harvard University, 1956. (ditto)

Henry, J., & Spiro, M. E. Psychological techniques: Projective tests in field work. In A. L. Kroeber (Ed.), *Anthropology today.* Chicago: Univ. of Chicago Press, 1953, Pp. 417–429.

Hodges, J. L., Jr. Discriminatory analysis 1. Survey of discriminatory analysis. Report No. 1, Contract No. AF41 (128)-8, University of California, 1955.

Kumata, H. A factor analytic investigation of the generality of semantic structure across two selected cultures. Unpublished doctoral dissertation, University of Illinois, 1957.

Kumata, H., & Schramm, W. A pilot study of cross-cultural methodology. *Publ. Opin. Quart.*, 1956, 20, 229–238.

Osgood, C. E., Suci, G. J., & Tannenbaum. P. H. *The measurement of meaning.* Urbana, Ill.: Univ. of Ill. Press, 1957.

Suci, G. J. A comparison of semantic structures in American Southwest culture groups. *J. abnorm. soc. Psychol.*, 1960, 61, 25–30.

Triandis, H. C., & Osgood, C. E. A comparative factorial analysis of semantic structures in monolingual Greek and American college students. *J. abnorm. soc. Psychol.*, 1958, 57, 187–196.

Vogt, E. Z. Navaho veterans, A study of changing values. Papers of the Peabody Museum of Harvard University, 41, No. 1, 1951.

(Manuscript received August 1, 1960.)

Much is unknowable.
No problem shall be faced
Until the problem is.
I, born to fog, to waste,
Walk through hypothesis,
An individual.

THOM GUNN, "Human Condition,"
The Sense of Movement

Crudely, there are three reasons models are built. First, models are either normative or descriptive. Within the class of descriptive models, there are norm-descriptive and case-descriptive ones. Case-descriptive models, unlike the other kinds, are not generalizations. They represent a set of referent causal processes by detailed analysis of one, or a few, specific cases. Then such questions arise as: Is the model sufficient? Does it represent the real situation? How can one test its value? For case-descriptive models, the usual likelihood tests are not appropriate because they are designed to evaluate generalizations. Outlined here, in addition to a frame of reference for the building of case-descriptive models, are suggestions for testing them.

TESTING CASE-DESCRIPTIVE MODELS

*by William H. Starbuck**

Purdue University

INTRODUCTION

DURING recent years, the general availability of large computers has fostered the development of a new kind of model-building. A small, but growing, number of researchers are constructing models which attempt to describe their referent causal processes in much greater detail than has been feasible heretofore. The ethic of these model-builders, to the extent that a common ethic exists, seems to be description for description's sake. The potential normative applications of their models are at best a secondary consideration. Similarly, the generality of their models is a secondary goal, because they believe that the most effective way to achieve efficient and accurate generalizations is through a detailed and elaborate analysis of the individual case.

The purpose of this paper is to outline a frame of reference for the builder of this new kind of model and to suggest a procedure for testing the models which are built.[1]

CASE-DESCRIPTIVE MODELS

A model is an attempt to represent a set of causal processes observed in the "real world." The set of causal processes being represented constitutes a *referent situation*. Therefore, a *model* is an operator which generates a set of variables X, given a set of variables Y, such that (a) the sets of variables X and Y represent characteristics of a referent situation, and (b) the operator itself represents the causal processes relating the characteristics of the referent situation.

A test is a critical examination of the model within the context of the model-builder's purpose. There seem to be three crude categories of models with respect to purpose. First, one can distinguish between *normative* and *descriptive* (or positive) models. Normative models are constructed with application in mind. In the purest case, the normative model is designed to answer a few predetermined questions. The ethical bases of the posed questions determine (a) which characteristics of the referent situation should be included in the model, and (b) how accurately these characteristics should be represented. Descriptive models are constructed for the purpose of describing and analyzing their referent situations. The model-builder frequently does not pose specific questions before constructing his model; or if he does pose questions beforehand, these questions have little ethical content. As a result, the builder of descriptive models is relatively unconstrained in his selection of characteristics of the referent situation to be included in the model and

* This paper was written while the author was a Ford Foundation Fellow at the Graduate School of Industrial Administration, Carnegie Institute of Technology.

[1] For another discussion of a related but slightly different problem, see Cohen and Cyert (1961).

has no clear criterion for the accuracy with which these characteristics should be represented.[2]

Second, within the class of descriptive models, one can distinguish between *norm-descriptive* models and *case-descriptive* models. Norm-descriptive models are intended to be generalizations. They focus upon an aggregate or an average, rather than an individual case. Frequently, one finds that norm-descriptive models are intended to be abstractions as well as generalizations: they are designed to assess the value of hypothetical, ideal relations between aggregated characteristics of the referent situations. In general, the builder of norm-descriptive models is as much concerned with the simplicity and generality of his model as he is with its accuracy, although this is frequently an advantageous trade, the statistics of aggregation being what they are.

The case descriptive models are the relatively new class of models which was described in the introduction. They are intended to be detailed analyses of one, or a few, specific cases. The model-builder seeks to represent the referent causal process as accurately as possible. Case-descriptive models do not normally attempt to be abstractions, except in so far as every model is an abstraction. In fact, adherence to a specific piece of unaggregated data is not sympathetic with an objective of discovering ideal relationships. This does not mean that the builder of case-descriptive models is not interested in generalization. He often believes that the most effective key to generalization — whether generalizations about individual cases or generalizations about an aggregate—is the detailed analysis of a specific case.

These three categories of models are neither exhaustive nor clearly defined. They do, however, serve as a useful factoring against which the model-builder can measure his own objectives, and thereby help him to decide what kinds of tests are most appropriate to his model.

[2] A more elaborate discussion of the distinction between normative and descriptive models can be found in Friedman (1953, Part I).

Examples may make these three categories clearer. Consider the following problem: human subjects are asked to predict which of two possible events will occur next in a series. The series of events is independent of the subjects' predictions.

A *normative* problem would be specification of the predictive strategy which the subjects should use. An appropriate model would represent the series of events and have just one output variable, the next event. If the model predicts correctly a large proportion of the time, it can be considered a good representation of the event series. Often, in normative problems, a specific degree of accuracy is stated as part of the objective, or can be derived from cost data. One of the simplest models of the event series is a biased coin. If the bias were known, the optimum strategy would be "predict on every trial the event which occurs most frequently in the event series as a whole." However, because the bias is not known, the strategy must be changed to "predict on trial N the event which has occurred most frequently in the previous $N - 1$ trials." If N is odd and both events have occurred equally often, the dilemma can be resolved by flipping an unbiased coin. A standard for good performance can be inferred from the model itself, since the model implies a probability density function for the number of correct predictions.[3]

[3] Assume that the two events occur with known probabilities p and $1 - p$. The a priori probability, $Q(N; p)$, of predicting the Nth event correctly is

$$Q(2k; p) = Q(2k + 1; p)$$
$$= (1 - p) \cdot \sum_{y=0}^{k-1} b(y; 2k - 1, p) + p \cdot \sum_{y=k}^{2k-1} b(y; 2k - 1, p),$$

and the a priori probability of predicting R out of N events correctly is

$$Pr[R; N, p] = Q[N; p) \cdot Pr[R - 1; N - 1, p] + [1 - Q(N; p)] \cdot Pr[R; N - 1, p].$$

Since $Pr[0; 0, p] \equiv 1$

and

$$Pr[0; N, p] = [1 - Q(N; p)] \cdot Pr[0; N - 1, p],$$

$$\sum_{\lambda=0}^{\infty} Pr[\lambda; N, p] \cdot S^{\lambda} = \prod_{\lambda=1}^{N} [1 - Q(\lambda; p) + S \cdot Q(\lambda; p)]$$

at least for $| s | < 1$. Therefore $Pr[R; N, p]$ is the

A *descriptive* problem arising from the same basic situation would be analysis of the behavior of the human subjects. The model could be restricted to one output variable, the subjects' next prediction. However, there is no conclusive reason to bound the model at this point. The model could generate variables representing the subjects' emotional involvement or their verbal explanations for their predictions. The model-builder must try to explain any aspects of the situation which he believes contribute to understanding. The criterion for accuracy is even more unclear. Models must be judged on structure as well as output. Models which generate different numbers and kinds of outputs are not directly comparable. Frequently accuracy in generating one output must be sacrificed for accuracy in generating another output.

Two distinct approaches to this descriptive task have been made. A number of *norm-descriptive* models have been proposed (Bush & Mosteller, 1955, Ch. 13; Estes & Straughan, 1954). These models attempt to describe the aggregate behavior of several subjects who are presented with identical event series. They have some common characteristics: (1) They focus on one output variable, the Nth prediction. This variable is the most easily aggregated characteristic of the situation. (2) They are stochastic models. Stochastic treatment again reflects the aggregated nature of the output variable.[4] (3) The models are all consistent with the empirical observation that the proportion of subjects who predict a specific event is asymptotic to the frequency with which that event occurs.

At least one model-builder has attempted a *case-descriptive* model of the subjects' behavior. Julian Feldman (1959)[5] analyzed in detail the behavior of three specific subjects and proposed different models for the behavior of each. The outputs of the models include the subjects' stated reasons for their predictions as well as their predictions. Feldman's models are very complex since discrete changes were observed in the subjects' decision processes.[6] It would be difficult to aggregate these models by calculation, and practically impossible to aggregate the models analytically. Nevertheless, Feldman was able to take some qualitative steps toward a general model.

The need to test a model within the context of the model-builder's purpose is obvious. Descriptive models should not be expected to be as appropriate for answering specific questions as normative models which were constructed with these questions in mind. This is particularly evident with respect to parameter estimation for forecasting over a specific time horizon. Norm-descriptive models should not be expected to represent specific cases as accurately as the appropriate case-descriptive models. Case-descriptive models should not be expected to generalize to new referent situations as effectively as norm-descriptive models.

There is a substantial body of literature relevant to the testing of normative and norm-descriptive models. See, for example, Koopmans (1950) and Hood and Koopmans

sum, over all combinations of R integers out of N, of terms of the form

$$\prod_i Q(i;\, p) \cdot \prod_j (1 - Q(j;\, p)),$$

where i is an index over the particular R integers in a combination, and j is an index over the $N - R$ integers not in the combination ($j \neq i$).

Because p is not known, its value must be inferred from the event series. If the event corresponding to p has occurred x times in the preceding N trials, and if p and x have a priori uniform distributions, the a posteriori conditional distribution for p is

$$Pr[p;\, N,\, x] = (N + 1) \cdot b(x;\, N,\, p).$$

Therefore the a posteriori distribution for R correct predictions in N trials is

$$Pr[R;N,\, x] = (N + 1) \cdot \int_0^1 Pr[R;\, N,\, p] \cdot [b(x;\, N,\, p] \cdot dp.$$

[4] Some model-builders take a position like that of Bush and Mosteller: "We tend to believe that behavior is intrinsically probabilistic" (1955, p. 3). However, even a theorist who believes that behavior is basically deterministic, can accept a probabilistic model which is macroscopic.

[5] Feldman reviews the norm-descriptive models in some detail, and his discussion of the testing problem is especially relevant here (1959, p. 130–136). A second study by G. P. Clarkson is in progress. An initial report on Clarkson's study has appeared in Clarkson and Simon (1960).

[6] Subjects used distinctly different decision processes after correct predictions and after incorrect predictions. Moreover, discrete changes in the subjects' hypotheses about the structure of the event series were frequent.

(1953). There is, however, reason to question the relevance of this literature for the testing of case-descriptive models. The usual approach to model testing focuses upon the probable event. It asks, given the model, how likely is the set of observations. A "good" model is one with a high probability of generating the observations. The builder of case-descriptive models can reasonably point out that he really does not care how likely his observations are since (a) he does not pretend to have constructed a generalization, and (b) the rare event does happen. Because the represented decision processes are both discrete and complex, case-descriptive data more nearly represent one point in an N-dimensional space than N points in a one-dimensional space. Moreover, from a strictly pragmatic point of view, the functional forms which appear in case-descriptive models are much more complicated than those which have yielded to the likelihood treatment. The test procedure suggested below is both more consonant with the objectives of the builder of case-descriptive models and more amenable to computation than is the usual approach.

TESTING SIMULATIVE POWER

The purpose of a case-descriptive model is to represent the referent causal processes. The ideal test would be based on the accuracy with which the model represents these referent processes. However, causal processes are nearly always the subject of inference, not of observation. There is no standard for the structure of the referent causal processes per se which can serve as a criterion for the model's accuracy. Therefore it is necessary to adopt some surrogate criterion.

The obvious choice for a surrogate criterion is the simulative power of the model. A *simulation* is an operator which generates a set of variables X, such that (a) the set of variables X represent characteristics of a referent situation, and (b) the set of variables X is indistinguishable from the corresponding characteristics of the referent situation. The word indistinguishable is used loosely since (a) the variables in the model represent only measured characteristics of the referent situation, not the physical entities which comprise the referent situation, and (b) distinguishableness is a binary property only with respect to an explicit criterion.

Characteristics of the referent situation may be either *observable* or *unobservable*—this being a function largely of the existing technology of measurement. Characteristic A is observable if its value can be measured without assuming any causal relationships between characteristic A and other characteristics of the referent situation. Characteristic A is unobservable if (a) its value can only be measured in conjunction with the values of other characteristics, or (b) no measurement technique is available for measuring its value. Here, the value of a characteristic is the state of the characteristic with reference to any arbitrary class of states, and measurement is any empirical restriction on the value of the characteristic.

A simulation is a testable operator because it does not claim to represent unobservable characteristics of the referent situation. A simulation claims only to generate variables which are indistinguishable from specific corresponding characteristics of the referent situation. The subset of characteristics may be very small indeed, relative to the set of all characteristics of the referent situation.

In practice, no model takes full advantage of the available measurement techniques. The variables in the model represent *observed* and *unobserved* (latent) characteristics of the referent situation. Characteristic A is observed if it is observable and measurements of its value are available for tests of the model. Characteristic A is unobserved if it is unobservable or it is observable but no measurements are available.

At this point, it might be well to say something about the use of the word representation. A case-descriptive model is not an end in and of itself. Each model should be a foundation for other models which include additional observed characteristics of the referent situation and which accomodate more precise measurements of the values of characteristics. Some unobserved variables, while not observed in the strictest sense, may be partially observed. By in-

corporating these partially observed characteristics of the referent situation, the model becomes a more effective foundation for other models and, at the same time, a more credible statement about the referent situation. Any frequent participant in academic seminars will note that the degree of acceptance of the model being presented, and by implication its effectiveness as a representation of the referent situation, depends in large part on the extent to which the model incorporates (or at least is consistent with) partially observed variables.

This is one point at which Friedman's position breaks down.[7] As long as models are constructed without regard for their consistency either with other related models or with later refinements of the same models, one simulation is as good as another. Nevertheless, most model-builders like to feel that their models *represent* something, and recognize at the same time that theirs is not likely to be the last word said on the subject.

The archetype for the case-descriptive model is, then, a *simulative model* which is testable as a simulation and at the same time represents the causal processes of the referent situation. Formally, a simulative model is an operator which generates a set of variables X, given a set of variables Y, such that (a) the sets of variables X and Y represent characteristics of the referent situation, (b) the set of variables X is indistinguishable with respect to an explicit criterion from the corresponding characteristics of the referent situation, and (c) the operator itself represents the causal processes relating the characteristics of the referent situation.

Being process models, case-descriptive models are necessarily dynamic. The generated set of variables X and the measurements of the corresponding characteristics of the referent situation are time series. In this paper, it will be assumed that the time series are discrete. The vector of measurements during a given time period, T, will be called an observation and denoted by an asterisk. $X(T)$ is the vector of generated variables at time T and $X^*(T)$ is the vector of measurements at time T.

It is frequently argued that no examination of a model is a critical examination, a true test, unless the test is based on a set of observations of the referent situation different from the set of observations used in developing the model. This may be a valuable comment on the testing of normative or norm-descriptive models; it is a much less valuable comment on the testing of case-descriptive models. First, rarely if ever does a model-builder pretend to have explained all aspects of the referent situation. Nearly all models purport to be just what they are: defective contributions to the over-all progress of human learning. Second, the effort invested in gathering the detailed data necessary to support a case-descriptive model is high relative to the work invested in the model itself. If the model is not consistent with the second set of observations, as in all probability it is not, it is better to spend time revising the model than to spend time criticizing the original model. Third, the value of the predictive criterion is that it exposes characteristics of the referent situation which were not previously observed; but since the model-builder's picture of the world is faulty, it does this at the risk of introducing more unobserved characteristics. As a result, the predictive criterion is more appropriate for testing models which aim at generality than it is for descriptive models. The same basic objective of exposing unobserved characteristics can be achieved by including in the model as many observable characteristics of the referent situation as possible.

TESTING MICROPROCESSES

The causal processes in the referent situation generate a set of observed variables X_0^* and a set of unobserved variables U_0, given a set of observed variables Y_0^* and a set of unobserved variables V_0. In a formal sense, the set of observed variables X_0^* defines the referent situation.

The model of the referent situation defined by X_0^* will be called the *macroprocess*. The macroprocess is a set of causal functions

[7] The essence of Professor Friedman's position is ". . . The only relevant test of the validity of a hypothesis is comparison of its predictions with experience" (1953, p. 8-9).

which generates a set of observed variables X_0 and a set of unobserved variables U_0, given a set of observed variables Y_0. For the time being, the sets X_0 and U_0 will be called endogenous variables and the set Y_0 will be called exogenous variables, although this terminology will be modified.

There is no set of unobserved exogenous variables V_0 in the specification of the model. Logically, this is a statement about the kinds of data which must be collected. Practically, this means that any characteristic of the referent situation which is not represented in the macroprocess or in the set Y_0 must either (a) have negligible causal effect upon the set X_0, or (b) be constant over the time span being simulated.[8] This same restriction is not true of the set, U_0, of unobserved variables which are represented in the model.

In general, the macroprocess can be factored into more than one mutually independent *microprocesses*. Formally, a microprocess is similar to the macroprocess. The Kth microprocess is a set of causal functions which generates a set of observed variables X_K and a set of unobserved variables U_K, given a set of observed variables Y_K. X_K is the set of observed variables endogenous to the Kth microprocess; it is composed only of elements of the set X_0. U_K is the set of unobserved variables endogenous to the Kth microprocess; it is composed only of elements of the set U_0. Y_K is the set of observed variables exogenous to the Kth microprocess; it is composed of both (a) elements of the set Y_0, and (b) elements of the set X_0.

One important characteristic distinguishes a microprocess from an arbitrary subset of the functions in the macroprocess. A microprocess is defined such that the functions which comprise it involve no unobserved variables except those which are endogenous to that microprocess. The only variables linking microprocesses are observed variables. The value of this characteristic is that a microprocess can be tested separately from other elements of the macroprocess. Tests of case-descriptive models should not be only binary judgments of the goodness of the model. Tests should make a constructive contribution by locating the sources of error in the model. By testing the microprocesses independently, sources of error are associated with specific subsets of the macroprocess. However, it is not sufficient to test only the individual microprocesses. A small error which does not appear significant when only one microprocess is tested may be cumulative when two or more microprocesses are combined (Cohen, 1960). The optimal testing procedure would sequentially examine all combinations of n microprocesses for $n = 1, 2, 3, \cdots$ until finally the macroprocess is tested as a unit.

The exogenous-endogenous terminology has been used rather loosely up to this point. Endogenous variables are variables which are generated by the model. In principle, these variables may be causally dependent on any characteristic of the referent situation. Exogenous variables are variables which are not generated by the model; they appear in the model because they have some causal influence on the endogenous variables. There is a causal arrow which points from Y_0 to X_0. For the sufficiency test outlined below, this is the only constraint placed on the selection of exogenous variables. However, if the descriptive model is to be used normatively, exogenous variables must be selected more carefully. The character of the causal arrow from X_0 to Y_0 becomes important. A particular exogenous variable $Y_i(T)$ may be causally independent of the sets $X_0(T)$, $X_0(T-1)$, $X_0(T-2)$, $\cdots$ $X_0(T-L_i+1)$ but causally dependent on $X_0(T-L_i)$, $X_0(T-L_i-1)$, $X_0(T-L_i-2)$, $\cdots$ $X_0(-\infty)$. This feedback from the referent causal process to its environment can radically affect the time series of a normative projection. Therefore, it is meaningful to speak of the *order* of the model, the order being the largest number of time periods for which the model can be validly projected. In the notation above, the order of the model is the smallest element in the set of causal lags L_i which characterize the causal relations from X_0 to Y_0. For example, if the set X_0 is projected for n time periods

[8] This point is discussed more fully in Simon and Guetzkow (1955, Appendix).

into the future, the order of the model must be n or greater; the causal functions defining Y_0 must not involve the variables $X_0(T)$, $X_0(T-1)$, $X_0(T-2)$, $\cdots$ $X_0(T-n+1)$.

TESTS OF SUFFICIENCY

The set X_0 is a probabilistic representation of the set X_0^*. It is a probabilistic representation for four reasons. First, measurements of the characteristics of the referent situations are inaccurate. Rarely is the set of feasible states of a referent characteristic known exactly and rarely can a referent characteristic be placed unambiguously in a defined state. Frequently measurements are crude; the feasible states of a referent characteristic are poorly defined and the state of the characteristic is ambiguous. An example of a characteristic that is accurately measurable is the state of a flipped coin: heads or tails. An example of a characteristic that is not accurately measurable, at least at present, is the amount of co-operative interaction during a unit time period of group activity. Second, most models contain *stochastic generators*, elements whose state is randomly determined. For example, the model may generate variable X, with either causal function R_{11} or causal function R_{12}, the choice between R_{11} and R_{12} being random. Third, the model-builder may make outright errors in specifying the causal functions. And fourth, it can be assumed that nearly all causal functions specified in the model are abstractions of the referent causal processes.

The first two sources of uncertainty will be called *loci of ignorance*. Loci of ignorance are aspects of the referent situation which (a) the model-builder accepts as unknown, and (b) have *ex ante* measurable effects on the simulated variables. In principle, inaccuracies in measurement and stochastic generators have measurable effects on the set X_0. The dispersions of the variables X_0^* and Y_0^* may not be known exactly but can usually be estimated, either by making multiple observations in each time period or by using a favorably designed measurement technique. The dispersions of stochastic generators are, of course, a similar case.

The second two sources of uncertainty are not loci of ignorance. First, the effects of abstraction or an error are not *ex ante* measurable. A major reason for building models is to evaluate the uncertainty attached to the errors and abstractions in the model. Second, the builder of case-descriptive models does not accept an abstraction as a desirable end in and of itself, only as a necessary truncation of his analysis of the referent causal processes. Outright errors are undesirable in any model.

The ideal objective of descriptive model-building is a set of functions which are both necessary and sufficient to explain the referent situation. There is, of course, no general test for this. A test of necessity implies that there is a meaningful characterization of all possible models and measurements on enough observable variables to make a unique choice among the possible models. A test of sufficiency, in the most general sense, implies that there is a meaningful characterization of all possible observations on all observable variables.

It is possible to test the sufficiency of a model, however, (a) with respect to a specific set of observed endogenous variables, (b) over a specific sequence of observations, and (c) given the loci of ignorance specified in the model. The set of observed endogenous variables, X_0, define the purpose of the model as a simulative device. The sequence of observations $X_0^*(T)$, $X_0^*(T+1)$, $\cdots$ $X_0^*(T+m)$ defines, in a practical sense, the referent situation. The loci of ignorance define the maximum accuracy with which X_0 should be generated given the detail to which the model has been extended. Thus, the test will focus on the uncertainty introduced by the sources not quantified as loci of ignorance: abstraction and outright error.

The greater the uncertainty attached to the predictions of the model, the less can be said about the accuracy of the model. The statement that a model is sufficient is a statement that the model is an accurate description of the referent causal processes. If large loci of ignorance are present in the model, the model is more likely to be accepted as sufficient but is less likely to be

an accurate description of the referent situation. Progress in description is made by reducing the sizes of loci of ignorance through more measurements, more accurate measurements, and more detailed analyses of the processes represented by stochastic generators. Through this process, models are subjected to increasingly rigorous tests of their accuracy.

Gradually, the sufficiency tests become more rigorous. More variables are observed; more observations are accumulated; the sizes of loci of ignorance are reduced. The set of sufficient models becomes smaller and smaller. In the long run, sufficiency converges toward necessity.

A case-descriptive model is, however, more than sufficient as defined by the sufficiency test. The role of the model as a representation of the referent causal processes—by incorporating in the model partially observed characteristics of the referent situation—is a weak guarantee that the model will be consistent with later models. That is, measurement of previously unobserved variables is unlikely to lead to drastic revision of the model.

THE SUFFICIENCY TEST

The set of unobserved endogenous variables U_0 can be factored into two subsets: (a) the set of unobserved endogenous variables which change over time U_0', and (b) the set of unobserved endogenous variables which do not change over time U_0''. The set U_0' is defined in terms of other variables by the causal functions of the model, as is the set X_0. Thus, the total number of causal functions in a microprocess is the sum of the number of variables in U_K' and the number of variables in X_K. The variables in the set U_0'' are free parameters; they are not determined by measurement and may exist only in the context of the specific model. However, logical limits may be imposed on these free parameters a priori.

The sufficiency criterion which is proposed here can be stated: "Does a set U_K'' exist such that the model has a probability greater than zero of generating the sequence of observations?" For each time period a subset, $P_k(T)$, of the possible free parameter values will have probability greater than zero of generating $X_K^*(T)$, given all lagged values of X_K^*, and all current and lagged values of Y_K^*.[9] The intersection, P_K^*, of $P_K(T)$ over all T defines sufficiency. If P_K^* contains one or more points, the microprocess is sufficient. If P_K^* is null, the microprocess is not sufficient.

The test considers only the *ranges* of X_0^*, Y_0^*, and the stochastic generators. In order for the test to be meaningful, the density functions of the loci of ignorance must be bounded. The model-builder cannot use density functions which have probability greater than zero for any value of the random variable (e.g., the normal distribution).

The sufficiency test can be criticized as being both too tight and too loose. It is tight in the sense that the model must be able to generate every X_0^* for the entire sequence of observations. In comparison to tests usually given theoretical models, this is fairly restrictive, but the builder of descriptive models should exchange something for the freedom to make specific rather than general models. The test is loose in the sense that only the ranges of the loci of ignorance are considered, not specific values. However, bounded density functions must be used, and the improbable event occurs.

This approach to testing suggests that the model-builder must, in the process of gathering data, focus his attention upon measures of the ranges of random variables rather than upon specific values. This is a deviation from usual practice, which tends to ignore measurements of the loci of ignorance except in so far as these are inferred from the fit of the model. In fact, in the construction of case-descriptive models, it is as important to measure the loci of ignorance as it is to measure the variables X_0^* and Y_0^*.

Example. As an example of the mapping of the subset $P_K(T)$, consider the causal function:

$$X \leftarrow AY + B$$

[9] The regions $P_K(T)$ must be bounded by the logical limits imposed a priori on U_K''.

where A and B are parameters with logical limits:

$$0 \leqq A$$
$$0 \leqq B$$

At time period T, the measurements of X and Y indicate that:

$$0 \leqq X_0 \leqq X \leqq X_1$$
$$0 \leqq Y_0 \leqq Y \leqq Y_1$$

Therefore the region $P_K(T)$ is the region in A-B space defined by the inequalities.

$$A \geqq 0$$
$$B \geqq 0$$
$$B \geqq X_0 - AY_1$$
$$B \leqq X_1 - AY_0$$

SUMMARY

This paper has attempted to outline a methodological ethic appropriate to the construction of case-descriptive models. The essence of the argument is that, since the referent situations for case-descriptive models are essentially samples of size one, tests should focus upon the sufficiency of the models using probability-not-zero as a criterion. This approach has computational as well as ethical advantages over the usual maximum likelihood criterion. The sufficiency test does, however, mean that the data-gathering process must focus more explicitly upon loci of ignorance than would ordinarily be the case.

Since small errors tend to accumulate in large, complex models, it was suggested that the macroprocess be factored into microprocesses which can be tested sequentially.

REFERENCES

Bush, R. R., & Mosteller, F. *Stochastic models for learning.* New York: Wiley, 1955.

Clarkson, G. P., & Simon, H. A. Simulation of individual and group behavior. *Amer. Econ. Rev.*, 1960, 50, 920–932.

Cohen, K. J. *Computer models of the shoe leather, hide sequence.* Englewood Cliffs, N. J.: Prentice-Hall, 1960.

Cohn, K. J., & Cyert, R. M. Computer models in dynamic economics. *Quart. J. Econ.*, 1961, 75, 112–127.

Estes, W. K., & Straughan, J. H. Analysis of a verbal conditioning situation in terms of statistical theory. *J. exp. Psychol.*, 1954, 47, 225–234.

Feldman, J. An analysis of predictive behavior in a two-choice situation. Unpublished doctoral dissertation, Carnegie Institute of Technology, 1959.

Friedman, M. *Essays in positive economics.* Chicago: Univ. of Chicago Press, 1953.

Hood, W. C., & Koopmans, T. C. (Eds.) *Studies in econometric method.* Cowles Foundation Monograph 14. New York: Wiley, 1953.

Koopmans, T. C. (Ed.) *Statistical inference in dynamic economic models.* Cowles Foundation Monograph 10. New York: Wiley, 1950.

Simon, H. A., & Guetzkow, H. Mechanisms involved in group pressures on deviate members. *Psychol. Rev.*, 1955, 62, 56–68.

(Manuscript received June 2, 1960.)

All I am or ever hope to be is due to irritations of my surface, together with such latent tendencies to response as may have been present in my original germ plasm.

W. V. QUINE, "The Scope and Language of Science"
British Journal for the Philosophy of Science

Although the value of diagnosis in physical medicine is well established, psychiatrists and psychologists, impatient with vague descriptive labels of undemonstrated value, decry the effort spent in diagnosing patients. While this author agrees with them in the belief that attaching a vague diagnostic label seldom helps the patient, he feels that the value of systematic diagnosis of schizophrenics never has been tested adequately. This paper does not take sides in the "dynamic" versus diagnostic controversy but attempts to formulate one approach to careful testing of the value of diagnosis in the area of schizophrenia.

MEDICAL AND STATISTICAL-PSYCHOLOGICAL MODELS FOR RESEARCH IN SCHIZOPHRENIA

by Wilson H. Guertin

University of Florida

THE desire to bring schizophrenia into the realm of clearly defined disease entities is a noble one; but, unfortunately, investigators engaged in the task resemble dogs chasing their own tails more than they do scientists. Whether the approach be psychodynamic, physiological, or psychological, there has been little added to the early classification of Kraeplin.

It is generally agreed that schizophrenia is such an inclusive label that there is no single behavior symptom always present in every patient, nor does a group of schizophrenics show much uniformity in traits or symptoms. For example, investigation of a trait such as hostility in a heterogeneous sample of schizophrenics will reveal about the same average amount as in a control group of normals. This, despite the fact that paranoids are very hostile and some other types are low in hostility. When these subsamples are combined, important features of each are lost and the over-all sample gives "normal" values although many of the individual subjects possess the characteristic to an abnormal degree.

While most investigators recognize and avoid such pitfalls, unfortunately the best way to avoid unproductive research in schizophrenia is never to begin. A few investigators make "careful" studies of schizophrenia and come out with statistically significant findings, which others are later unable to confirm. This is not surprising since there is no quantitative, or even reasonably satisfactory, way to describe the nature of the original sample of subjects so that others can duplicate the experimental conditions. It seems clear that a more exact systematic classification or description of schizophrenics is necessary for facilitating research and furthering psychiatric knowledge.

MEDICAL MODEL

Psychiatrists discussing classification often distinguish between syndromes and disease entities (Noyes, 1940, p. 161). A syndrome is merely a "concurrence of symptoms" which may or may not reflect a specific etiology and pathology; whereas a disease entity is formulated in terms of etiology and pathology.

Syndromes represent earliest attempts to order facts about clinical cases in a convenient way when etiological and pathological features remain obscure. Kraeplin described the medical model for disease very clearly: "Judging from our experience in internal medicine, it is a fair assumption that similar disease processes will produce identical symptom pictures, identical pathological anatomy, and an identical etiology. If, therefore, we possessed a comprehensive knowledge of any one of these three fields—pathological anatomy, symptomatology, or

etiology—we would at once have a uniform and standard classification of mental diseases. A similar comprehensive knowledge of either of the other two fields would give not only just as uniform and standard classifications, but all these classifications would exactly coincide. Cases of mental disease originating in the same cause must also present the same symptoms and the same pathological findings" (Diefendorf, 1921, p. 117).

An example of the application of the medical model in internal medicine might be cited, such as the development of an understanding of jaundice. First would come the observation that many patients show both yellow skin and sclerae. An examination shows considerable variability of secondary symptoms. Some have sudden and others slow onsets; also, some have a light-colored stool that persists throughout the jaundice while others do not. The intercomparison of onset and stool color discloses that the light stool goes with the acute onset (obstructive versus infectious). Thus two distinct subtypes are identified on the symptom level. Further study of atypical cases reveals a group of new-born infants (erythroblastosis fetalis) as well as a group of adults (post transfusion) with some common secondary symptoms despite the age difference. Then, laboratory and autopsy investigations attack all of the syndrome groups separately. Distinct pathological anatomy and etiology are eventually established for each subtype and the result is a sound disease classification for the jaundices. Further work concentrates on breaking down hepatic infection on the basis of the presence of viruses or bacteria.

Psychiatry has not progressed past the first stage of labeling gross syndromes because of the numerous interdependent behavior variables present, many of which are irrelevant to the syndrome but hard to ignore. Or, putting it more simply, the personality is too complex to be understood through gross observations. It is at this point that the psychological statistician must help in the identification of symptom subtypes to bring about systematic order in the search for underlying diseases. The psychologist is trained in techniques of observing and recording behavior, while the statistician knows how to intercompare symptoms and analyze the cases to yield basic symptom types.

STATISTICAL-PSYCHOLOGICAL MODEL

The most unique characteristic of psychology as a science is the study of individual behavior. While other characteristics of schizophrenics such as physiology, anatomy, socio-environmental, and historical background variables could be employed for the first systematic classification of subtypes of schizophrenics, the present model begins with behavioral characteristics or symptoms. It is the psychologist's bias that the study of behavior variables will be the most profitable starting point.

The statistical aspects of the model grow from a need to examine behavioral similarities of patients objectively. It is this similarity in patients, when repeatedly encountered, that suggests the presence of a distinct subtype. Correlation coefficients are a convenient way of expressing the degree of similarity or difference between patients. Thus, the first statistical step in the development of the model is to provide for intercorrelating a number of schizophrenics possessing heterogeneous behavioral characteristics. The individuals themselves lie on the phenomenal or concrete level of the diagram in Figure 1, while ratings of their behaviors are shown on the first level of abstraction. Intercorrelations of the subjects according to obtained ratings provide the second level of abstraction, while clusters of intercorrelations of similar subjects occur at the third level. The last and final level, the proposed types, based upon factor analysis, will be discussed at length after a statistical digression.

In any analysis of intercorrelations, one must first decide whether he wishes to end up with clusters of people or clusters of symptoms. We are committed to look for clusters of similar individuals because the medical model seeks more than a concurrence of symptoms, viz., ultimately identifying a specific disease in a specific individual. Thus, a transposed matrix of intercorrelations is in

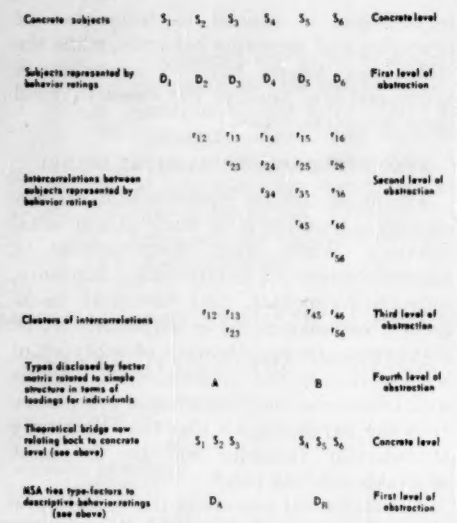

Concrete subjects
S_1 S_2 S_3 S_4 S_5 S_6
Concrete level
Subjects represented by behavior ratings
D_1 D_2 D_3 D_4 D_5 D_6
First level of abstraction
r_{12} r_{13} r_{14} r_{15} r_{16}
r_{23} r_{24} r_{25} r_{26}
r_{34} r_{35} r_{36}
r_{45} r_{46}
r_{56}
Intercorrelations between subjects represented by behavior ratings
Second level of abstraction
Clusters of intercorrelations
r_{12} r_{13} r_{23}
r_{45} r_{46} r_{56}
Third level of abstraction
Types disclosed by factor matrix rotated to simple structure in terms of loadings for individuals
A
B
Fourth level of abstraction
Theoretical bridge now relating back to concrete level (see above)
S_1 S_2 S_3
S_4 S_5 S_6
Concrete level
MSA ties type-factors to descriptive behavior ratings (see above)
D_A
D_B
First level of abstraction

FIG. 1. Conceptual Scheme for Subtyping Individuals According to Correlational Similarity

order and we intercorrelate people on the basis of relevant symptoms.

McQuitty (1959) proposes a process of discovering clusters of similar people called linkage-analysis. The method is simple but does not provide a unique solution since the derived clusters depend largely on arbitrary judgements made by the classifier. In effect, it is a cluster analysis technique and like other cluster analysis procedures it can be a first step in analyzing intercorrelations, but eventually we must turn to factor analysis and the unique solution provided by the simple structure concept (Cattell, 1952, p. 145).

It should be recalled that in the process of intercorrelating people, all information about symptoms disappears and derived types are defined in terms of subjects. This makes it awkward for the investigator trying to report his findings publicly since he cannot convey meaningful information by saying that "Type *A* is exemplified by heavily loaded subjects 1, 2, and 13." The characteristics of these subjects are unknown to the reader so the type must be described by some reference back to symptoms. While transposed analysis holds the only present promise of providing unique typological solutions, this descriptive shortcoming has served to limit its application. Because of this difficulty, conventional factor analyses often are used inappropriately in attacks on typological problems since it is so much more convenient to have the symptoms themselves enter the factor table and give implicit meaning to the derived factors.

It is true that symptom data merely disappear during intercorrelation and are recoverable. Thus, when the investigator describes the nature of the individuals with heavy loadings on a factor, he can mention the common symptoms they possess. The relationship between a symptom and type-factor can be represented quantitatively by a correlation coefficient or chi-square but neither quantitative method is entirely satisfactory. Product moment correlations will be low if only individuals with extreme factor loadings show marked deviation on the behavior item rated and there is little covariation for subjects lying in the intermediate range of factor loadings. Unsuspected curvilinear relations tend to reduce these coefficients as well. Chi-squares can reveal the association of dichotomously scored symptoms with a particular range of factor scores but only when the cutting-line is placed appropriately.

The technique of Manifest Structure Analysis (MSA), introduced by du Mas (1956), can be viewed as systematically exploring the placement of cutting-lines to disclose associations between symptom categories and particular ranges of scale values such as type-factor loadings. As portrayed in Figure 1, factor analysis leads to types that are defined in terms of subjects. These types are redefined in terms of the abstracted symptom behavior data by MSA and this step becomes an integral part of the statistical-psychological model (see Figure 1).

Up to this point we have been concerned with the first stage in establishing schizophrenic disease entities, that of identifying phenomenal groupings of schizophrenics on the basis of symptom similarities. After these gross types have been established,

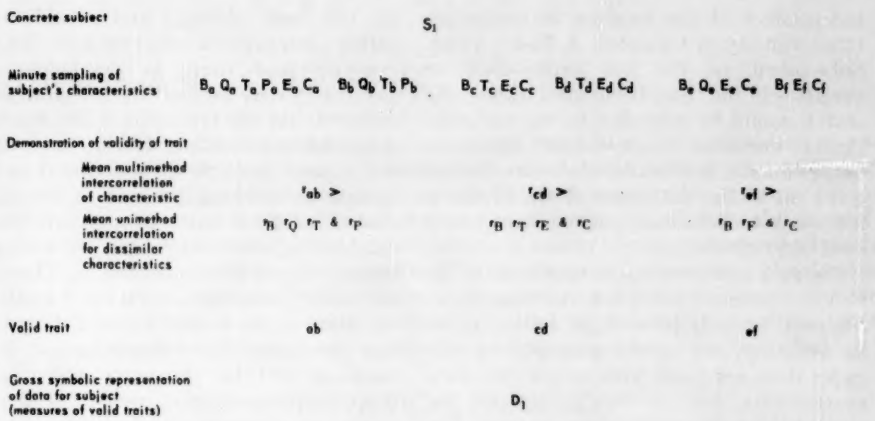

Concrete subject
S_1
Minute sampling of subject's characteristics
$B_a Q_a T_a P_a E_a C_a$ $B_b Q_b T_b P_b$ $B_c T_c E_c C_c$ $B_d T_d E_d C_d$ $B_e Q_e E_e C_e$ $B_f E_f C_f$
Demonstration of validity of trait
Mean multimethod intercorrelation of characteristic
$r_{ab} >$ $r_{cd} >$ $r_{ef} >$
Mean unimethod intercorrelation for dissimilar characteristics
$r_B r_Q r_T$ & r_P $r_B r_T r_E$ & r_C $r_B r_F$ & r_C
Valid trait
ab cd ef
Gross symbolic representation of data for subject (measures of valid traits)
D_1

FIG. 2. Expanded View of the Characteristics of a Single Individual as They are Related to the Characteristics of Other Individuals

they must be purified and cross-validated to assure optimum reliability and inclusiveness of classification. Irrelevant symptoms must be discarded during the first stage when it becomes obvious that they are leading to typological bases other than schizophrenic pathology. For example, symptoms related to sex difference could subdivide a single type into two small clusters. Conventional factor analysis of the behavior symptoms is highly suitable for identifying these trouble-making variables that may cluster enough to establish some irrelevant typological bases.

MSA will cause most unreliable and irrelevant symptom items to be excluded from further typological consideration. Next it becomes necessary to add new items and recast discarded ones by referring to those that have proved helpful in establishing the preliminary types. MSA item analysis provides a quick means of detecting which of the added items should be retained. After old and new items have been combined, a new transposed factor analysis on a second sample of subjects will disclose the amount of improvement in specifying types (reliability) and illustrate the inclusiveness of the types (invariance of factor structure).

Even at this stage, the number of behavior characteristics and the way they are sampled are quite limited when compared with all possible types of variables. As yet, no use has been made of biological and environmental variables, which offer greatly expanded possibilities for sampling relevant features of schizophrenics. Also, psychological tests and questionnaires offer fruitful sources for variables. The psychologist's bias and skills would encourage him to look first to the other behavior samples (tests and questionnaires) before introducing the biological and environmental variables.

The investigator would look for stable traits that would be represented on all behavior levels and search for representative measures at each of these levels. For example, the original behavior rating of impulsiveness (assuming it is a trait) should have its counterparts in a test where impulsiveness leads to a high error score and in a questionnaire item acknowledging the tendency to act on the spur of the moment. True, some types might be too hostile or guarded to co-operate on questionnaires, but the missing data on some levels of behavior present no problem.

It would appear desirable to keep each class of variables separate in the early search for types in order to hold down the total number of variables, and to provide for later cross-validation of the original types with other classes of variables so that it could be demonstrated that they are

independent of the medium of evaluation (trait validity of Campbell & Fiske, 1959). Subsequent to the full exploitation of variables in the three behavioral realms, the search would be extended to the variables in the domains of physiology, anatomy, home climate, and course of disease. Before going on to the last phase of the model, a few words to the statistically oriented would seem appropriate.

Figure 2 represents an enlargement of the first two levels of the schematic diagram in Figure 1 but only for a single individual, *S*. In actuality, the model proposed in this paper does not work with simple raw data abstractions, but utilizes traits as the variables for intercorrelating people. The word traits is used in the statistical sense discussed by Campbell and Fiske (1959). A trait is a construct based upon "valid" measurement by more than one "method."

S_1 in Figure 2 stands for subject number 1. His characteristics are studied by many methods as shown—*B*ehavior Rating, *Q*uestionnaire, Psychological *T*est, *P*hysiological Measurement, *E*nvironment, and *C*ourse of Illness. One characteristic (subscript *a*) is measured by all methods while another, *b*, may not permit measurement through observation of environment or course of illness. This segment of the model constitutes the multitrait-multimethod matrix of Campbell and Fiske, and if there is a high intercorrelation of all these multimethod measures and low cross-correlation with the other clusters of measures, then the trait is established.

Conventional factor analysis can play an important role here in establishing the validity and nature of the traits in the multimethod design of Figure 2. Ideally, the factor analytic establishment of traits should precede the transposed analytic search for types. The type search would then depend upon the intercorrelation of people on the basis of valid and reliable traits rather than isolated behavior data. Revered concepts and traits in psychology and psychiatry may be used for intercorrelating subjects, but the procedure is not nearly as dependable as using factor analytically established traits. Still, the tremendous task of first establishing traits is so formidable that it is difficult to envision an investigator establishing traits and still proceeding to identify types all in one lifetime. Therefore, the model presented in Figure 1 omits the requirement for formally establishing traits and presumes that the behavior rating variables will be chosen to reflect some degree of abstraction in terms of traits.

The last phase in completing the medical model is the examination of physiological, anatomical, etiological, and course of illness characteristics of each of the established symptom types. As a study of these areas consistently reveals etiological and pathological constancies for a given type, the disease entity itself finally will be established. New specific treatment methods may be expected to develop coincident with such investigation.

REFERENCES

Campbell, D. T., & Fiske, D. W. Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychol. Bull.*, 1959, 56, 81–105.

Cattell, R. B. *Factor analysis.* New York: Harper, 1952.

Diefendorf, A. R. *Clinical psychiatry.* New York: Macmillan, 1921.

du Mas, F. M. *Manifest structure analysis.* Missoula: Montana State Univ. Press, 1956.

McQuitty, L. L. Differential validity in some pattern analytic methods. In B. M. Bass & I. A. Berg (Eds.), *Objective approaches to personality assessment.* Princeton: Van Nostrand, 1959. Pp. 66–82.

Noyes, A. P. *Modern clinical psychiatry.* Philadelphia: W. B. Saunders, 1940.

(Manuscript received January 19, 1961.)

CRITIQUE AND COMMENT

All political theories make assumptions about man and society, and therefore all political scientists must be to some extent both psychologists and sociologists. Modern political science has had considerable difficulty in encompassing the insights and discoveries of modern psychology and in relating private motivations to public actions. A survey of current research suggests that most studies have failed to integrate these two dimensions of behavior. Here an analysis of Erik Erikson's theories of ego identity suggests that the dynamics of personality development tend to emphasize precisely those theories which are basic to any effectively communicated political ideology.

PERSONAL IDENTITY AND POLITICAL IDEOLOGY*

by Lucian W. Pye

Center for International Studies, Massachusetts Institute of Technology

I

MUCH of political science can be thought of as an attempt to understand the connections between individual and group behavior. Such a statement may not provide the most satisfactory beginning for a definition of the discipline, but it does accentuate the concern of political science with both the nature of the individual and the nature of society. Political theorists have traditionally been interested in the state and the statesman, the law and the law-maker, community and the citizen, historical forces and individual choices. All political theories must be premised on either explicit or implicit assumptions about the nature of man and of society; hence the political scientist must be to some extent both a psychologist and a sociologist.

These were the qualities of the traditional political scientists. The most important forerunners of the discipline were also in their time the leading authorities on questions about human nature and the structure of society. Thus, Aristotle and Machiavelli, Hobbes, Locke, and Rousseau were concerned not just with the traits and arts of political leaders, but with the basic nature of man; and in Plato we have some remarkable anticipations of modern psychology. Similarly, traditional theorists were concerned with more than just the polity; their interests usually covered the entire web of human associations. They were remarkably successful in building their systematic political theories out of the most advanced psychological and sociological knowledge available to them, producing a synthesis which related their views about the character of man and society to normative standards and judgments about the design of public institutions. For example, one strand of this Western tradition of political science can be readily summed up under the rubric of reason, an essentially psychological concept. The universe was conceived as rational and man as having the rational power to understand it; given the right sort of institutions men would act rationally according to their enlightened self-interest, and the result would be a happy and harmonious social order.[1] Other systems, of course, were erected on the premise that man's essential nature is brutal, selfish, and inherently evil. Clearly the traditional political theorist did not shirk from psychological considerations; on

* Written for *Political Decision Makers*, edited by Dwaine Marvick, Copyright 1960, by the Free Press, Glencoe, Illinois.

[1] For an excellent outline of the psychological premises of classical political theory and of the problems posed by modern psychology, see Cook (1959).

the contrary, he often permitted his psychological notions to color all aspects of his theories.

The contemporary political scientist wishing to follow in the tradition is confronted with an infinitely more difficult task. The first problem is the extraordinary rate of growth of specialized knowledge about psychological and sociological matters. Many political scientists have come to feel that it is hopeless to attempt to incorporate these intellectual developments into systematic political theories; and therefore, finding it impossible to follow the practices of the earlier political theorists, they have decided instead to examine systematically, and with reverence, the works of these earlier thinkers. Those who have sought to continue in the tradition have found a second and more fundamental problem in our new understanding of the basic nature of man which has come from Freud's discoveries. In providing insights into the full dimensions of man's inner nature, Freud made it embarrassingly clear that previous theories about political relationships were generally premised on impoverished and inadequate notions of human nature. The concept of rational man has had to be altered. At the same time Freud held out the promise of a deeper understanding of political phenomena, a promise that has not been readily realized.

It is one thing to suggest that all political theories depend upon some psychological view of man, and quite another matter to relate the insights of psychoanalysis to political analysis. In spite of the enthusiasm of those who are ready to try to enrich political science with Freud's psychoanalytic contributions, it must be acknowledged that the results are often awkward, and at times even grotesque. In what meaningful ways can the political scientist benefit from the contributions of those who, in dealing with our disturbed citizens, have altered our image of man?

In performing the pathbreaking role of applying psychoanalytical concepts to political analysis, Harold D. Lasswell (1930a) placed the relationship between private motivations and public acts at the center of his approach. He suggested that the dynamics of political action were to be found in the configuration of the individual personality and not just in the grand issues of history. Political man is characterized as being moved by private motives which are displaced on public objects and rationalized as being in the national interest. By employing the assumption about the nature of man that dominated psychoanalytical thinking during the 1930's, Lasswell was led to the view that release of tension is fundamental to all political behavior, and the key to human action in general. We find him writing that: "Nations, classes, tribes, and churches have been treated as collective symbols in the name of which the individual may indulge his elementary urges for supreme power, for omniscience, for amorality, for security" (1930b, p. 39). And again he stated that: "Indeed one of the principal functions of symbols of remote objects, like nations and classes, is to serve as targets for the relief of many of the tensions which might discharge disastrously in face-to-face relationships (1930b, p. 73. Quoted in Lynd, 1958, pp. 97–98).

Lasswell's efforts to trace the connections between private motivations and public acts represented an attempt to deal, on a systematic basis, with the kinds of problems that political biographers have long sought to untangle: the relationship between the child and the man, between personal peculiarities and political preferences, between private frustrations and public ambitions. Lasswell's critics have charged that he exaggerated the importance of private and irrational considerations because he so undervalued the significance of political issues and rational choice. On the other hand, Lasswell's critics have generally failed to appreciate that, despite the special attention he gave to psychological matters, he steadfastly maintained a remarkably rigid distinction between the psychological and the political, between the pattern of personality development and the logic of institutional change, between private motivations and public policy.

Thus, in advancing the view that the social process consists of people pursuing *values*, through *institutions* by means of

available *resources*, Lasswell (1948, pp. 16–17) implied that the individual and his value preferences are one thing, and social institutions another. Some institutions may be more appropriate than others for maximizing particular values, but in the last analysis institutions are based on functions that involve more than individual preferences. Similarly, Lasswell pointed to the likelihood of certain personality types being more successful in particular political roles, but he rigidly held that the basic character of any political role is determined by its functions in the political process and not for the personality. By posing the problem in these terms, Lasswell avoided the error, frequently made by psychologists, of seeing the *homo politicus* as a distinct personality type characterized by an inordinate craving for power. He defined instead a variety of political roles and political personality types (1948, Ch. II–IV). In particular, Lasswell identified particular personality types with the political roles of the administrator, the agitator, and the theorist (1930a, pp. 53–55).

Lasswell's method of bringing psychoanalytical considerations to bear on political analysis opened the way to a tremendous growth in empirical studies of factors traditionally considered to be "nonpolitical." (Indeed, some critics have held that "political behavior" studies are nothing more than studies of the nonpolitical, i.e., the social and personal aspects of group behavior.) Lasswell's formulation of the relationship between personality and politics encouraged, for example, studies of the social backgrounds of political elites, the symbols of political identifications, and the informal factors that influence decisions.[2] The same assumptions about the relationships between personality and politics can be found in most of the studies of voting behavior (Gosnell, 1942; Lazarsfeld, Berelson, & Gaudet, 1948; Berelson, Lazarsfeld, & McPhee, 1954; Campbell, Gurin, & Miller, 1954). Implicit in most such studies is the notion that personality and "informal" considerations are largely "irrational" with respect to the logic of public institutions.

There have been numerous attempts to bridge the gap that Lasswell has left between the private and the public, between the dimension of personality and the sphere of politics. In general these attempts may be divided into two categories. First, there are those efforts to find direct correlations between the formation of specific personality types and political behavior. These are attempts to see the political from the perspective of psychological insights, and they are generally concentrated on the dynamics of personality formation during childhood.[3] The second category consists of those attempts to fill the gap between the psychological and the political with other social processes. Instead of seeking to move directly from personality formation to political action, the psychological factor is related to all the other factors which might influence political behavior. The "psychological" dimension thus includes cognitive processes. In the main, this approach has centered on the study of various aspects of the political socialization process.

Most studies in the first category are contributed by cultural anthropologists and others involved in work on personality-and-culture. In conceiving of culture as the pattern of basic values reflected in all phases of life, the anthropologist has assumed a very close and direct relationship between personality and political behavior. This approach has largely taken the form of national character studies; the names of Ralph Linton, Margaret Mead, Ruth Benedict, and Geoffrey Gorer, come readily to mind. Historians have, of course, long employed unspecified notions about the characteristics of different nations and peoples.[4] Psychoanalytically oriented students of national character have been much more explicit, and consequently, it has been easy for their

[2] The most outstanding examples would be the Hoover Institute Studies in the RADIR project which were authored by Daniel Lerner, Ithiel deSola Pool, Robert North, and others.

[3] Among the best statements of the methodology of this form of psycho-cultural analysis are: Leites (1948), Kardiner (1945), and Mead (1951).

[4] The advantages for historians of the newer insights of psychoanalysis are well stated by William L. Langer (1958).

critics to take exception to particular details.[5]

Within the category of attempts to relate directly personality and politics are those which seek to establish a connection between specific personality configurations and susceptibility to particular political ideological orientations. The outstanding example of this form of study is, of course, *The Authoritarian Personality* (Adorno, Frenkel-Brunswik, Levinson, & Sanford, 1950). This monumental study, originally conceived as a search for the possible psychological sources of anti-Semitism, became an attempt to demonstrate congruence between personality type and political ideology. Despite its impressive reception in the social sciences, most students of political behavior have found it difficult to accept its suggestion of a direct correlation between authoritarian personality types and membership in authoritarian political movements.[6]

A characteristic common to such studies is that the treatment of psychological matters has been methodologically considerably more sophisticated than the political analysis. Subtlety in psychological insights has not necessarily yielded significant knowledge about the political realm. Advances that stem from public opinion research have further encouraged study of intervening social processes that serve to reduce the gap between personality formation and political behavior. These studies, which we have classified in the second category, tend to be more influenced by orientations common to social psychology than by psychoanalysis. The shift has also been away from a central interest in childhood development and toward an appreciation of the total immediate social context within which action takes place.

For example, Gabriel Almond in his study of the appeals of Communism in four European countries (1959) found it necessary to trace the total process of personality development from the early determinants of personality to the later cognitive developments. By working in terms of political socialization, Almond developed the concept of "political culture." The process of becoming politically involved is seen as being similar to, but distinct from, the process by which an individual becomes a member of his culture. Out of the early childhood experiences, and later influenced by the processes by which he is introduced to the political world, the individual is finally recruited into a specific political role (Almond, 1956, 1960). Subjective psychological factors can be treated at all stages in the analysis while, at the same time, objective factors relating to the social and political setting can be given their full weight. Other attempts to narrow the gap between personality and politics are to be found in the works of Smith, Bruner, and White (1956); Hyman (1959); Katz and Sarnoff (see Sarnoff & Katz, 1954; Katz, McClintock, & Sarnoff, 1957; Katz, Sarnoff, & McClintock, 1956); and Isaacs (1958). The rate of advances in political psychology is impressively demonstrated by the recent books of Lane (1959) and Lipset (1960) which systematically build upon the best of current research in the field.

There are two principal reasons why these various authors have been able to add the psychological dimensions to their studies without grossly oversimplifying basic political considerations. First, they have balanced their analysis of the shaping of the unconscious with explicit treatment of ego-functioning in the total development of the personality. Second, they have related personality factors to the social context of action and, wherever possible, to specific sociological variables. The link between "personality" and political behavior is established by recognizing that personality development depends upon the individual's perceptions of social reality, his emotionally conditioned responses to his environment, and his learned modes of evaluating reality. This approach has led some scholars to a rigorous search for relationships between attitudes, opinions, and "basic orientations"

[5] David Potter, an historian, has written an excellent and very sympathetic critique of much of the work on national character. See Potter (1954).

[6] Herbert McClosky (1958) has continued to search for correlations between political conservatism and a set personality type.

on the one hand, and social or demographic distinctions on the other. Others, like Erich Fromm (1941) and David Riesman (1950), have related more broadly defined personality or character types to the main social or economic configurations of an historical period. A highly original and complex type of analysis has been Daniel Lerner's treatment of the connections between basic aspects of personality and the dynamics of social change in transitional societies (1958).

The use of this image of personality, largely derived from social psychology, has been extremely rewarding when combined with survey techniques. It focuses attention on the relationship between "personality type," basic social and political attitudes, and the distribution of each according to demographic and socio-economic categories.[7] But there is some concern among those who have worked with survey techniques that studies built largely upon a psychology of the cognitive processes will give too "flat" a picture, lacking in nuances that depend upon the functions of the unconscious. In order to avoid this problem it may become increasingly necessary to supplement social survey studies with "interviews in depth." These approaches all represent significant advances in the utilization of psychological theory for political analysis. It would seem, however, that we have yet to arrive at satisfactory methods of exploiting for political studies the full potentialities of psychoanalytical theory. Every now and then there is the rare work, such as that of Nathan Leites (1953),[8] which taps this potential and demonstrates what can be realized if the new image of man is truly incorporated in a political analysis. Thus although much has been accomplished in recent years to enrich political science with the insights of modern psychology, we are still seeking new ways of bridging the gap between the dynamics of individual behavior and the developmental forces of history.

[7] An outstanding example of such a study is one by Inkeles and Bauer (1959).

[8] For two excellent commentaries on Leites' methods and his analytical assumptions, see Bell (1958) and Kluckhohn (1955).

II

It is from such a background of experiences as we have just outlined that the political scientist looks for guidance to Erik H. Erikson's *Young Man Luther: A Study in Psychoanalysis and History* (1958). In what ways can the clinical analyst, who has already made great contributions to ego psychology and to cultural anthropology, assist the student of history?

Erikson is concerned in this volume with the problem of the great man in history, who creates the ideologies in which people find their beliefs. He is concerned to clarify how the lone individual, in seeking to find himself and give meaning to his own character, can give shape and form to a period of history. More precisely, Erikson is interested in the relationship between the identity crisis in the personality development of the reformer and the ideology he creates. Erikson reduces the gap between the public and the private spheres, between the psychological and the political, by stressing the connecting links between personal identity and public ideology.

For some time it has not been fashionable for political scientists to grapple with the problem of the great man who tries to change the course of history. Modern political science grew up during a time when the notion of "science" was largely one of a quest for uniformities. This assumption encouraged an orientation that found its highest expression in sociological and economic analyses. Even when the pendulum began to swing back to a greater emphasis upon the individual and, hence, upon psychological considerations, attention still centered largely on aggregates and not on the unique individual. For the search has been mainly for relationships between social and economic factors and personality types. Recent interest in the study of decision-making and of elite roles has centered mainly on the webs of relationships within which individual decision-makers have to fight their way, and not upon the function of the leader's personality. One exception to this tendency that comes readily to mind is Alexander and Juliette George's provocative

analysis of Woodrow Wilson's personality (1956).

III

In turning to the content of Erikson's analysis we must begin with his key concept of the identity crisis. This is the crisis of late adolescence when the young person, after having synthesized and resynthesized the experiences and reactions of each of the earlier stages of childhood, must move out of childhood and assume a place in the adult world.[9] Erikson has sought significance in all the typical characteristics of this stage of life: the periods of moodiness and sentimentality; the restless spirit but lethargic body; the sense of ambition and the desire to explore and know all possibilities, but also the endless moping and hanging around; and the unexpected vacillations between excessive worldliness and unbelievable naïvete—between trying to be more adult than adults, and then being more childish than children. Above all it is the crisis of uncertainty when the youth must commit himself—usually after many fitful starts—to a definition of himself that he and others will recognize. Erikson says, "I have called the major crisis of adolescence the *identity crisis*; it occurs in that period of the life cycle when each youth must forge for himself some central perspective and direction, some working unity, out of the effective remnants of his childhood and the hopes of his anticipated adulthood; he must detect some meaningful resemblance between what he has come to see in himself and what his sharpened awareness tells him others judge and expect him to be."

Erikson has emphasized elsewhere that the term identity expresses a ". . . mutual relationship in that it connotes both a persistent sameness within oneself (self-sameness) and a persistent sharing of some kind of essential character with others" (1956a, p. 57). Thus *identity formation* "arises from the selective repudiation and mutual assimilation of childhood identifications, and their absorption in a new configuration, which, in turn, is dependent on the process by which a *society* (often through subsocieties) *identifies the young individual* recognizing him as somebody who had to become the way he is, and who, being the way he is, is taken for granted. The community, often not without some initial mistrust, gives such recognition with a (more or less institutionalized) display of surprise and pleasure in making the acquaintance of a newly emerging individual. For the community, in turn, feels recognized by the individual who cares to ask for recognition; it can, by the same token, feel deeply—and vengefully—rejected by the individual who does not seem to care" (1956a, pp. 68–69).

It becomes immediately apparent that Erikson's concept of ego identity provides a far more complex and multifaceted model of the human personality than is customarily used in political analysis. His concept covers not only the individual's constitutional givens and his idiosyncratic libidinal needs, but also the nature of his cognitive processes and the historically specific quantities of information which he has stored in his memory. The concept goes beyond this for it implies that there are systematic relationships among not just these dimensions of the personality, but also the particular mental or physical faculty favored by the individual and best developed, his effective psychological defense mechanism, his successful sublimations and even the degree of consistency with which circumstance has required him to assume particular roles.

The political scientist is not one to pass judgments on the technical aspects of Erikson's concept of ego identity.[10] In order to appreciate, however, the ways in which Erikson relates the great ideological reformer to his times and to history, it is necessary to understand the outlines of Erikson's theories about the development of the individual. This is because he assumes

[9] Erikson has developed his theories of ego-identity in numerous writings: *Childhood and Society* (1950); "Ego Development and Historical Change" (1946); "Growth and Crises in the 'Healthy Personality' " (1953); "On the Sense of Inner Identity" (1953); "The Problem of Ego Identity" (1956a).

[10] The political scientist with curiosity about ego psychology may want to read in addition to Erikson's writings: Anna Freud (1952), H. Hartmann (1951), and H. Hartmann (1950).

that the way in which the major problems of each consecutive stage in the development of the individual are met will be reflected in the evolving configuration of that individual's identity; and in the case of the great man this means that his peculiar pattern of development will be reflected in his political ideology. Erikson thus implies that there is a minimum but fundamental structure to any ideology which is related to elemental aspects of personality development. Therefore, in reviewing Erikson's theories about the development of the individual, we shall focus on those aspects that are most relevant in contributing to the tone and the spirit of ideologies.

It is only in the last chapter of Erikson's book on Luther that he sets forth his theory of the various stages of personality development. Indeed, those readers who have not previously been exposed to his views on the subject may find that they can get more out of his analysis of Luther if they read this chapter as an introduction. For our purpose of summarizing the stages of development and suggesting their relationship to the identity crisis, we may rely upon a diagram Erikson has devised (1956a, p. 75). It should be noted, however, that a diagram often suggests well-defined and rigorous relationships without actually articulating their precise nature. This is an important warning, and as Erikson has remarked, such a diagram "can be recommended to the serious attention only of those who can take it *and* leave it" (1956a, p. 76).

The diagram is so constructed that the diagonal line of squares reading from the upper left hand corner to the lower right present the major crisis in personality development for each of the stages of growth which are listed in the left hand margin from top to bottom. The basic pattern of personality development from infancy, to early childhood, and on through to adolescence and the mature age is outlined along this diagonal. Each crisis is of a dialectical nature, and in the diagram what would generally be considered as the more successful outcome of each crisis is stated first.

The diagram also shows the relationship of the other stages to the critical stage of the identity crisis. The vertical column 5 states the ways in which the identity crisis is foreshadowed during each of the earlier crises. There is thus in each of these earlier ages some aspect of the identity crisis. The horizontal row V outlines the principal implications for the question of identity of the way in which the other crises may be resolved. Thus the horizontal row V should be matched up with the diagonal line of squares, while the vertical column 5 should be related to the age categories.

Thus if we begin with the beginning we have the crisis of *basic trust* during infancy which corresponds in large measure with Freud's oral stage. Out of this first social relationship between the mothering adult and the mothered child, the individual either gains that first and most fundamental of all psychosocial traits, that original "optimism," that assumption that "somebody is there," that treasure of "basic trust." Or, denied the necessary security, he comes to a profound sense of mistrust which will color his entire existence; he will miss something others can take for granted most of the time, and, according to Erikson, only psychiatrists, priests, and born philosophers can appreciate how sorely he will miss it (1958, p. 118). The question of identity at this age takes the paradoxical form of the "positive" development being the lack of differentiation or a sense of "unipolarity," while the "negative" development would be a premature sense of self-differentiation. Erikson's concept of a sense of unipolarity seems to go beyond the more traditional Freudian concept of narcissistic omnipotence which is customarily associated with this age. The infant not only feels that he can command his world, but more fundamentally he feels that the world is for him and he can be for the world. If on the other hand, the infant feels that in his first relationship he is unrelated to the other, that the world can ignore him and he can have no control over it, then the consequence will be a sense of isolation, of premature self-differentiation, of basic mistrust.

Shifting next to the subsequent consequences for identity formation of the basic trust crisis (1, V), the main issue according

	1	2	3	4	5	6	7	8
I Infancy	Trust vs. Mistrust				Unipolarity vs. Premature self-differentiation			
II Early Childhood		Autonomy vs. Shame, Doubt			Bipolarity vs. Autism			
III Play Age			Initiative vs. Guilt		Play identification vs. (Oedipal) Fantasy-Identities			
IV School Age				Industry vs. Inferiority	Work Identification vs. Identity Foreclosure			
V Adolescence	Time Perspective vs. Time Diffusion	Self-Certainty vs. Identity Consciousness	Role Experimentation vs. Negative Identity	Anticipation of Achievement vs. Work Paralysis	Identity vs. Identity Diffusion	Sexual Identity vs. Bisexual Diffusion	Leadership Polarization vs. Authority Diffusion	Ideological Polarization vs. Diffusion of Ideals
VI Young Adult					Solidarity vs. Social Isolation	Intimacy vs. Isolation		
VII Adulthood							Generativity vs. Self-Absorption	
VIII Mature Age								Integrity vs. Disgust, Despair

to Erikson is that of the relationship of the individual to time. With trust comes a sense of time perspective, an optimism about the future, a feeling that good things will come with waiting, that stress will soon be relieved with pleasure. On the other hand, a failure of the ego function to maintain this perspective is related to an early inability to develop satisfactory expectations about need-tensions and their satisfactions. Time diffusion is a basic mistrusting of time: every delay appears as deceit, every need to wait becomes an experiencing of impatience, every hope a signal of danger, every potential provider a probable traitor.

To go beyond the level of individual identity to that of ideological formulations, the crisis of basic trust is related to the role of faith. With a perspective on time and a powerful sense of trust, Utopias become possible, planning and programming give hope and meaning, and loyalty is in itself rewarding. When the efficacy of faith must be denied, planning and policy must lose all meaning and deceit is seen in all actions. Presumably a political ideology might meet the needs of some who have not had the full measure of basic trust. But according to Erikson, "Of all the ideological systems, however, only religion restores the earliest sense of appeal to a Provider, A Providence. In the Judeo-Christian tradition, no prayer indicates this more clearly than 'The Lord make His Face to shine upon you and be gracious unto you. The Lord lift up His countenance upon you and give you peace' " (1958, p. 118). What was to Freud the oral stage has become with Erikson that which also makes the "face" so important in human affairs: "face-to-face relationships," "face the facts," "face the future," "face up to life," "let's face it."

The next age of early childhood, comparable to Freud's anal stage, brings the crisis of autonomy as against shame and doubt. From this stage comes that element in the sense of autonomy "... which can and does mean independence, but does and can also mean defiance, stubbornness, self-insistence" (1958, p. 122). What is basic faith in the earlier age becomes human will, "in its variations of will power and willfulness" (1958, p. 255); and what is basic mistrust in the oral stage becomes shame and doubt in the second stage. Shame is different from and quite as important as guilt in Erikson's view. Shame is "the loss of social innocence, the blushing awareness that one can 'lose face,' have 'too much cheek,' and suffer the wish to be invisible, to sink into the ground. Defiance obviously, is shame's dialectical opposite: and it makes sense that the willful exposure of the behind came to mean a defiant gesture of shamelessness . . ." (1958, p. 122). With respect to the later identity crises, the issue of autonomy versus shame will be resolved either in favor of a sense of self-certainty, a sense of autonomy with will and purpose, or with a crushing sense of self-consciousness and a defiant need to claim a self.

It is impossible to summarize accurately Erikson's extremely subtle interpretation of young Luther's experiences during these first two crises of his life. It must be recognized that we are doing Erikson an injustice when we reduce his analysis to the bare statements that: first, Luther obtained an extraordinary reservoir of basic faith from his mother, and thus always had a deep understanding of the dynamics of faith. However, "Martin was driven early out of the trust stage, out from under his mother's skirts by a jealously ambitious father who tried to make him precociously independent from women, and sober and reliable in his work. Hans succeeded but not without storing in the boy violent doubts of the father's justification and sincerity; a lifelong shame over the persisting gap between his own precocious conscience and his actual inner state; and a deep nostalgia for a situation of infantile trust. His theological solution—spiritual return to a faith which is there before all doubt, combined with a political submission to those who by necessity must wield the sword of secular law—seems to fit perfectly his personal need for compromise" (1958, pp. 255–256). These are the origins, as Erikson sees it, of that combination of faith and wrath, of the belief that behind a God that shames and a God that is demanding there is still the

possibility of mutual recognition, of the face-to-face meeting of personal salvation. As Erikson puts it, "It would be much too easy (although some stalwart opponents of all interpretation would consider even this easiest and most obvious explanation far fetched) that Han's son was seeking in religion what he would not find in Hans" (1958, p. 115).

We have gone into detail with these first two stages in order to give a feeling for Erikson's mode of analysis. We must treat the other stages in more summary fashion.

The third crisis, that of initiative versus guilt, covers more than the Freudian concept of the Oedipus complex. Erikson reflects his personal values, and above all, his ability to be hard-headed and yet sympathetic, without being precious and sentimental, when he writes of the Oedipus complex. For example, he says, ". . . we . . . most certainly . . . would ascribe to Luther an Oedipus complex, and not a trivial one at that. We would not wish to see any boy—much less an imaginative and forceful one—face the struggles of his youth and manhood without having experienced as a child the love and the hate which are encompassed in this complex: love for the maternal person who awakens his senses and his sensuality with her ministrations, and deep and angry rivalry with the male possessor of this maternal person. We would also wish him with their help to succeed, in his boyhood, in turning resolutely away from the protection of women to assume the fearless initiative of men" (1958, p. 73).

Clearly, Erikson does not minimize the Oedipus complex, and yet it is not a dominant theme in his total analysis. The influence of the initiative versus guilt crisis on the subsequent identity crisis depends upon whether the individual resolves his Oedipus complex by turning outside of the family to seek his future development, and thus envisioning the possibility of experimenting with other roles. The alternative would be a sense of *negative identity*, that is, a need to become what one has been warned not to become, which is something one can only do with a divided heart. Since there seems to be a need to protect one's wholeheartedness those who have a sense of negative identity cannot even be steadfast rebels.

Erikson's analysis moves without interruption from this stage on into the crisis of industry and inferiority. In this fourth stage the individual's "budding will to phantasy, play, games and early work" are all related to the occupational and technological ideals which the child perceives in his environment. What for Freud was the phallic stage becomes with Erikson a period of systematic learning and of collaboration with others. "The resolution of this stage decides much of the ratio between a sense of industry or work completion, and a sense of tool-inferiority, and prepares a man for the essential ingredients of the ethos as well as the rationale of his technology . . . In Martin's case, the tool was literacy, Latin literacy, and . . . he was molded by it . . . and later he remolded, with the help of printing, his nation's literary habits. With a vengeance he could claim to have taught German even to his enemies" (1958, pp. 258–259). A major dimension of personality formation in Erikson's view is the learning of a craft or a skill, and the development of proficiency in manipulating particular tools, instruments, or symbols.

Next comes the identity crisis, the principal one in Erikson's analysis, in terms of which his book on Luther is written. The last three crises—those of intimacy versus isolation, generativity versus self-absorption and integrity versus despair—are fairly self-evident and need no further elaboration here. We need only note that in Erikson's view the "integrity crisis, last in the lives of ordinary men, is a lifelong and chronic crisis in a *homo religious*. He is always older, or in early years suddenly becomes older, than his playmates or even his parents and teachers, and focuses in a precocious way on what it takes others a life-time to gain a mere inkling of: the question of how to escape corruption in living and how in death to give meaning to life" (1958, p. 261).

IV

It has been considerably easier to set forth Erikson's views on the "metabolism of generations" than it will be to state his

theory about the dynamics of the great man in history. In this volume Erikson is still experimenting with his theory, and thus he is extremely reluctant to use any direct propositional statements. We can only outline what seems to be the direction of his explorations.

The story of the great man, ideologically speaking, is the story of an individual striving to find his own identity. By following the peculiar logic of that struggle, he gives, without necessarily intending to, a sense of identity and meaning to a people at a particular juncture of history. The individual pattern of development may be extremely complicated, but Erikson suggests that there are certain uniformities. In particular, he makes comparative references to the lives of Freud, Darwin, G. B. Shaw, and Hitler.

The beginning point, and indeed the central theme, of the development of the great man is his need to settle a personal account on a large scale and in a grand context. This involves far more than the idea that public life provides the opportunity for reducing psychic tensions. In Erikson's view, the great man must have some score to settle with others and the score must be of such grand proportion that there is an appropriateness in seeking a public arena. The problem must be far more than just the sting of the Oedipus complex. Erikson elsewhere has commented that: "Whoever has suffered under and identified with a stern father, must become a stern father himself, or else find an entirely different means of moral strength, an equal measure of strength. Young Martin Luther's religious crisis is a transcendent example of the heights and depths of this problem" (1956b, p. 50).

This magnitude of the personal problem and the compelling need to resolve it seems to produce a deep sense of ambivalence: a feeling of sinfulness on the one hand, and a feeling of being chosen on the other. The great man may come to seek greatness, to believe in his own destiny, and yet he can be consumed with a fear of failure, a precocious fear which may arise at a very early age. Erikson indeed suggests that such men often fail as children, in that they are people with lost childhoods—people who place excessive demands upon themselves from a very early age. In their lives the identity crisis is likely to involve a conflict between a sense of allness and a feeling of nothingness. The fear of failure becomes a dread of nonexistence, and the individual vacillates between a sense of nothingness and a sense of being everything. Erikson is convincing when discussing the thin line that separates the feelings of omnipotence and of insignificance.

A key element in the identity crisis (but something which may last through life) is the "moratorium." Erikson attaches considerable significance to the need of all individuals for a period of moratorium: a withdrawal from full involvement, a time of loneliness and uncertainty, a time during which psychic growth may catch up with physical development. Most societies have institutionalized in varying degrees this moratorium, in the sense that they withhold responsibility from young people who are physically as developed as adults. During the moratorium the young person generally develops some skill or technique which will either subsequently become central to his identity or give him the necessary sense of discipline so that he will be able to employ more effectively some other faculty or skill. Luther's moratorium was spent in the monastery; Freud's "monastery" was his medical and scientific training; G. B. Shaw spent his moratorium in a business house; Darwin spent his in medical training and two years aboard the *Beagle*; for Winston Churchill it was the Indian Army. Erikson suggests that the common pattern of the ideological innovator is one of coming upon his life work without prior planning or design. Disciplined training is in one area, and creative innovation in another.

To achieve his sense of identity and break out of his moratorium, the great man must rely upon some dominant faculty, some special gift. In Erikson's reading of history, there is also likely to be at this juncture some technological innovation which can peculiarly complement the special qualities of the great man and thus become the bridge in linking him to his times. In the

case of Luther, it was his understanding of the Word and Gutenberg's invention of the printing press. Erikson has some subtle things to say about playing with words, and about the importance people will attach to isolated words and characters, to doctrines and pronouncements, to the Word and to the Good Book. Even more important is Erikson's appreciation of compulsions that lie behind the need of the great ideological innovators to talk, to manipulate words. Above all else, such people do not need simply to talk: they need to *talk back*. It is not at the level of the content of their words, but at a much deeper level, that they really mean what they say. This is the need to settle their score; this is the real "Meaning of Meaning It."

Once the reformer has found his identity and commands a medium through which he can act without conflict, the stage is set for a great moment in history. This is the moment of consummation when both the reformer's search and the period's search for identities are mutually realized. Erikson suggests that this moment of success affects the ego of the leader and the ego of those reached by and receptive to his word. All who are involved know that they "mean it," and hence a historical force of great power is produced.

For the people a long-needed shift in outlook has occurred and they will never be the same again. For the leader this can only be a moment of balance and then the inner conflict must go on again. While Luther was "In the cloister, all three factors—his sense of identity, his potential for intimacy, and the discovery of his generative powers—were stubbornly engaged in the life-or-death struggle for that sense of total justification which both the father and The Father had denied him, and without which a *homo religious* has no identity at all" (1958, p. 149). Once he did find the main executive of his identity, the Word, he was soon caught up again in the conflict between a sense of nothingness and an urge for fame.

This conflict can result in the collapse of the innovator and even in his becoming in time the very thing he fought. But this is only part of the problem. The leader may also realize to himself that he really did not intend to do what he did on the grand scale; he may come to recognize that it was only a private score he had to settle. "The crisis of an ideological leader naturally emerges when he must recognize what his rebellion—which began with the application of a more or less disciplined phantasy to the political world in the widest sense—has done to the imagination, the sense of reality, and the conscience of the masses. The fact is that all walks of life, revolutionized but essentially leaderless, exploited Luther's reformation in all directions at once. They refused to let him, and a few people like him, settle down in parsonages as the representatives of the praying orientation in life, and otherwise accept the estate and occupation in which, as he claimed, God had placed them. The princes became more absolute, the middle-class more mercantile, the lower classes more mystical and revolutionary; and the universal reign of faith envisaged in Luther's early teachings turned into an intolerant and cruel Bible-quoting bigotry such as history had never seen. As Tawney put it, 'the rage of Luther . . . was sharpened by embarrassment at what seemed to him a hideous parody of truths which were both sacred and his own' " (1958, p. 242).

The final result is that the ideological leader produces consequences that are quite the opposite of those he intended. "Consider for a brief moment certain great names of our time, who pride themselves on a dominant identity enhanced by scientific truth. Darwin, Einstein, and Freud—omitting Marx, who was a conscious and deliberate ideological craftsman—would certainly deny that they had any intention of influencing, say, the editorials, or the vocabulary, or the scrupulosity of our time in the ways in which they undoubtedly did and do. They could, in fact, refute the bulk of the concepts popularly ascribed to them, or vaguely and anonymously derived from them, as utterly foreign to their original ideas, their methodology, and their personal philosophy and conduct. Darwin did not intend to debase man to an animal; Einstein did not preach relativism; Freud was neither a philosophical pansexualist nor a moral

egotist. Freud pointed squarely to the psychohistorical problem involved when he said that the world apparently could not forgive him for having revised the image of man by demonstrating the dependence of man's will on unconscious motivation, just as Darwin had not been forgiven for demonstrating man's relationship to the animal world, or Copernicus for showing that our earth is off-center" (1958, p. 177–178).

In summary, the great man in seeking to settle his own inner problem strikes out against his environment. He is both destructive and constructive; indeed, without a capacity for destruction he cannot be truly constructive. It is through leveling blows against the existing system that the great man first wins his audience. His public must also feel some uncertainty over the times, some loss of identity. Then comes the role of charisma, the bases of a new mode of communication, and the beginnings of a new sense of identity for both the leader and the followers. The particular faculty and the particular technology provide the foundations for institutionalizing the relationship of charisma and for giving structure to the new sense of identity. Then comes the extremist reactions: first the leader overdoes it, and then the followers overdo it.

V

What is Erikson's purpose in this book? And what special value does the study have for political scientists? In making his point that the consequences of ideological leaders are often the opposite of their intended purposes, Erikson presents a remarkably insightful comparison of Luther and Freud. There is considerable significance in each of the points he makes about Luther, who came at the end of the age of absolute faith, and Freud, who came at the end of the age of reason. We are, however, particularly struck by the following:

> Both men endeavored to increase the margin of man's inner freedom by introspective means applied to the very center of his conflicts; and this to the end of increased individuality, sanity, and service to men. Luther, at the beginning of ruthless mercantilism in Church and commerce, counterpoised praying man to the philosophy and practice of meritorious works. Subsequently, his justification by faith was absorbed into the patterns of mercantilism, and eventually turned into a justification of commercialism by faith. Freud, at the beginning of unrestricted industrialization, offered another method of introspection, psychoanalysis. With it, he obviously warned against the mechanical socialization of men into effective but neurotic robots. It is equally obvious that his work is about to be used in furtherance of that which he warned against: the glorification of "adjustment." Thus both Doctor Luther and Doctor Freud, called great by their respective ages, have been and are apt to be resisted not only by their enemies, but also by friends who subscribe to their ideas but lack what Kierkegaard called a certain strenuousness of mental and moral effort (1958, p. 252).

Here and there throughout the book, Erikson criticizes the popularizers of Freudian analysis. At the surface level he is particularly critical of their soft-headed and sentimental qualities and of their lack of appreciation for the high costs and the self-discipline that must go into any form of human excellence. Erikson repeatedly reminds us of a point that some popularizers of psychoanalysis ignore, which is that Freud did not merely seek to help the individual adjust to his environment, but that he also demanded that the individual make the "environment" adapt to him. The difference is extremely important for political analysis. A peculiar set of biases, largely favoring static analysis and equilibrium concepts, is likely to predominate if it is assumed that the individual must either adjust to and be molded by "society" or revolt against society. A radically different outlook comes from picturing the individual as not only adapting to others, but also in turn changing and controlling his surroundings. With such an orientation, initiative, creativity, and the basic dynamics for the system do not come from anything as abstract as "society" but rather from particular individuals. All people are to some extent creative, for all people must shape their worlds in finding their identities.

At a deeper level, Erikson's concern over what he considers to be the popular misinterpretations of Freud involves another problem which is left largely unarticulated, but which colors much of the analysis. This is the problem of the proper relationship of

the disciple to his ideological father. How can one find one's own identity while remaining true to one's leader and mentor? Here, of course, is the problem that provides the dynamic element behind all controversies about who is perpetuating the true faith, and who is deviating from it. This has been the problem of those who took their inspiration from Luther and also of those who took theirs from Freud. Erikson is known as a Freudian and yet he has quite possibly done more original and innovational work than any of the neo-Freudians. (Freud only used the term "identity" once, and with a rather different meaning from that of Erikson's.) A clue to Erikson's method in becoming both the complete follower and the creative innovator is to be found in the following words that he once used to sum up his appraisal of Freud:

> Freud, before he went into medicine, wanted to become a lawyer and politician, a lawmaker, a *Gesetzgeber*. When, in 1938, he was exiled from his country, he carried under his arm a manuscript on Moses, the supreme law-giver of the people whose unique fate and whose unique gifts he had accepted as his own. With grim pride he had chosen the role of one who opens perspectives on fertile fields to be cultivated by others. As we look back to the beginnings of his work, and forward to its implications, we may well venture to say: Freud, the physician, in finding a method of healing himself in the very practice of emotional cure has given a new, a psychological rationale for man's laws. He has made the decisive step toward a true interpenetration of the psychological with the technological and the political in the human order (1956b, p. 62).

Erikson thus claims that Freud was centrally concerned with the larger order of human relationships and that such historical studies as this one of Luther should belong at the heart of the intellectual revolution sparked by Freud. The insights of psychoanalysis should not be treated as marginal ideas that may add novelty to otherwise completed studies; instead, Erikson would hold that these are insights which must be placed at the center of political analysis and become an integrated part of research. Erikson, however, is quick to recognize that this will be a difficult development because there is such widespread misconception about what Freud's contribution really was. This is why Erikson is so critical of the superficial popularizers of psychoanalysis. He also acknowledges that psychoanalysis has ". . . developed a kind of *originology* . . . a habit of thinking which reduces every human situation to an analogy with an earlier one, and most of all to that earliest, simplest, and most infantile precursor which is assumed to be its 'origin.' Psychoanalysis has tended to subordinate the later stages of life to those of childhood. It has lifted to the rank of a cosmology the undeniable fact that man's adulthood contains a persistent childishness: that vistas of the future always reflect the mirage of a missed past, that apparent progression can harbor partial regression, and firm accomplishment, hidden childish fulfillment. . . . We must grudgingly admit that even as we were trying to devise, with scientific determinism, a therapy for the few, we were led to promote an ethical disease among the many" (1956b, p. 18–19).

Erikson's answer to this problem is that Freud's great contribution was not certain formulas about personality development, but a *technique of observation*. Throughout his book Erikson says that if we are to understand, analyze, and appreciate those who shape history, we must really look at them, and observe them as full individuals. Our ability to see others clearly depends, however, upon our readiness to take a hard introspective view of ourselves. This, of course, is where the argument becomes delicate and even sticky. Erikson, however, does not imply that only those who have, under the guidance of others, taken such a long introspective view are capable of sensitive and complete political analysis. Rather he is warning us that we should be especially careful whenever we feel that we are being especially "honest," "objective," or "open-minded." We must recognize that some of our claims to being "objective" may be prompted by a desire to protect ourselves from an honest look at reality. We must be prepared, as Erikson says, "to relinquish the security of seemingly more objective methods" (1956b, p. 60). There are times, of course, when the problem for the individual social scientist must become even more complex. In the words of Erikson:

> An adult studying a child, an anthropologist studying a tribe, or a sociologist studying a riot sooner or later will be confronted with data of decisive importance for the welfare of those whom he is studying, while the strings of his own motivation will be touched, sometimes above and sometimes well below the threshold of awareness. He will not be able, for long, to escape the necessary conflict between his emotional participation in the observed events and the methodological rigor required to advance his field and human welfare. Thus, his studies will demand, in the long run, that he develop the ability to include in his observational field his human obligations, his methodological responsibilities, and his own motivations. In doing so, he will, in his own way, repeat that step in scientific conscience which Freud dared to make (1956 b, p. 60).

Erikson's basic criticism of the popularizers of psychoanalytical concepts is that they have given to the social scientist a set of phrases which have become mechanical and wooden formulas, and which in turn have become the devices which shield those who use them from the painful tasks of introspection and of honestly viewing reality. It is not surprising that intelligent, sensitive, and knowing people have found dubious value in studies which seek to make a direct connection between the nursery and world affairs, between crib and cabinet.

The primary contribution of Erikson's work for the political scientist is the reminder that Freud's main contribution was a technique of observation, and that observation is the key to the study of human relationships as it is the key to all sciences.

What about the particular theories and concepts that Erikson has developed? He has created a model of personality development which can be of great interest to the political scientist. It is, however, such a complicated and subtle one that it can probably only be used to full effect by its creator. At this level, Erikson's contribution is likely to be mainly one of making us more sensitive to a wide range of emotional nuances and of subjective relationships in human development and behavior.

Erikson's concept of identity can have a more specific impact on political research. It becomes an exciting and illuminating concept when applied, for example, to the problem of political development in the newly emergent countries. These are societies whose peoples, in spite of babbling slogans of nationalism, lack a sense of identity. When old forms and customs lose their binding, their sustaining, and their reassuring powers, the people must restlessly search for new personal identities and for a new sense of collective identity. The arena is prepared for the ideological reformer; for, in Erikson's terms, the great man of history. In this setting it is all too easy for the shallow charismatic leader to appear, for a moment, as a prophet. Those who are facile with words may have great appeal, for the people need the word to find a new way. It is a time in which words are fundamentally more important than actions. It is also a time in which words become cheap and action becomes impossible. The setting is right for the politically anxious to try out—possibly with enthusiasm, but certainly without true commitment—all manner of ideological forms. Before the nation can develop, leaders must emerge who have found integrity in their own quests for identity, and who can hence speak in terms that will bring meaning to other people's search for identity. The need is for that set of shared orientations which force a people to face reality steadfastly, and which makes it impossible for them to run from reality. This is the meaning Erikson gives to ideology.

Erikson's concept of identity thus suggests that the problem of consensus in transitional societies runs deeper than mere agreement over political forms and over the appropriate ends and means of political action; it involves the creation of an inner coherence of values, theories, and actions for the entire polity. The implication is that in underdeveloped countries there is a vicious circle at the subjective level that is more crucial to the problem of national development than the more manifest vicious circle of poverty, ill-health, and illiteracy. Those who articulate the hope that national identity can come from modernization cannot escape the depressing psychological fact that modernity, in their mind, has always been the monopoly of precisely those who were their former masters. If they hated

their colonial rulers then they cannot expect to find their identities by following the same path. If they did not hate their former rulers there might still be the problem of preferring dependency to autonomy which would confuse the quest for identity. National identity cannot be built upon doing less well what one's former master excelled at. Similar psychological roadblocks appear when identity is sought among the recorded but forgotten remnants of a distant history, for such a search becomes, psychologically, a constant reminder of national impotence in recent history. These are only some of the subjective problems which impede the solution of the more objective problems in the underdeveloped countries, and which cannot be resolved until a sense of national identity is achieved.

The readily apparent utility of Erikson's concept of identity to the problems of the underdeveloped countries suggests that it must have a much broader value. Specifically, his concept can give focus to all forms of study which are concerned with the development of the political actor. By stressing the individual's need to find coherence in his self-image, Erikson suggests that we should see the process of political socialization not as a series of random experiences, but rather as a trend in development in which there is always a coherence in the way in which people come to their political orientations. Our search should thus be for more than connections between isolated attitudes and opinions on the one hand, and particular demographic characteristics on the other; it should be for the more complete pattern, the total configuration, the full style of political actors.

By demonstrating that the gap between the private and the political can be effectively bridged by the relationship between personal identity and political ideology, Erikson helps the political scientist with more than the problem of the psychological aspects of behavior. He also offers an approach for studying ideologies themselves. For some time political scientists have known that the logic of ideologies is not encompassed by mere reason. There has, however, remained the problem of a systematic method for treating the nonrational components. Erikson, in suggesting that political ideologies are an expression of the total character of the human personality, has provided us with an analytical framework to understand the inner structure of ideologies. Psychoanalytical theory can be a guide for comprehending more than just individual patterns of behavior. Nathan Leites has shown how extremely rewarding it can be to examine an ideology—or in his phrase, an operational code—according to the insights of psychoanalysis. Leites chose to leave his theory implicit; Erikson has stated his theory in more detail. Both have worked at the creative stage in the development of a promising approach, and the next stage will have to be one of increased precision in the statement of propositions and increased rigor in empirical testing.

Erikson is able to challenge and stimulate the thinking of political scientists because he combines so effectively in himself the qualities of both the scientist and the artist. Like a great artist, he evokes in others an urge to imitate him. Also, like a great artist, he makes us sensitive to the wonders of what we had taken to be mere commonplaces. In referring to his own youth, Erikson once noted, "I was an artist then, which is a European euphemism for a young man with some talent, but nowhere to go" (1956b, p. 40). In going far since then, Erikson has not lost the genius of the true artist.

REFERENCES

Adorno, T. W., Frenkel-Brunswik, Elsa, Levinson, D. J., & Sanford, R. N. in collaboration with Aron, Betty, Levinson, Maria H., & Morrow, W. *The authoritarian personality.* New York: Harper, 1950. Sponsored by the American Jewish Committee, Social Studies Series, Publication No. III.

Almond, G. Comparative political systems. *J. Politics*, 1956, 18, 391–409.

Almond, G. *The appeals of communism.* Princeton, N. J.: Princeton Univ. Press, 1959.

Almond, G. Theoretical introduction. In G. Almond et al., *The politics of the developing areas.* Princeton, N. J.: Princeton Univ. Press, 1960.

Bell, D. Ten theories in search of reality. *World Politics*, 1958, 10, 327–365.

Berelson, B. R., Lazarsfeld, P. F., & McPhee, W. N. *Voting.* Chicago: Univ. of Chicago Press. 1954.

Campbell, A., Gurin, G., & Miller, W. E. *The voter decides.* Evanston, Ill.: Row, Peterson, 1954.

Cook, T. I. Democratic psychology and democratic world order. *World Politics*, 1949, 1, 553–564.

Erikson, E. H. Ego development and historical change. In *The psychoanalytic study of the child*, Vol. II. New York: International Univ. Press, 1946.

Erikson, E. H. *Childhood and society.* New York: W. W. Norton, 1950.

Erikson, E. H. Growth and crises in the "healthy personality." Fact-finding Committee, Mid-century White House Conference. New York: Josiah Macy Jr. Foundation, 1950. Also in C. Kluckhohn and H. Murray (Eds.), *Personality in nature, culture, and society.* New York: Knopf, 1953. (a)

Erikson, E. H. On the sense of inner identity. In *Health and human relations.* Report on the International Conference in Hiddesen, Germany, 1951. New York: The Blakiston Co., 1953. (b)

Erikson, E. H. The problem of ego identity. *J. Amer. Psychoanalytic Assoc.*, 1956, 4, 56–121. (a)

Erikson, E. H. The first psychoanalyst. *Yale Review*, Autumn, 1956. (b)

Erikson, E. H. *Young man Luther: A study in psychoanalysis and history.* New York: W. W. Norton, 1958.

Freud, Anna. *The ego and the mechanisms of defense.* New York: Basic Books, 1952.

Fromm, E. *Escape from freedom.* New York: Rinehart, 1941.

George, A. L., & George, Julliette L. *Woodrow Wilson and Colonel House.* New York: John Day Co., 1956.

Gosnell, H. F. *Grass roots politics: National voting behavior of typical states.* Washington, D. C.: American Council on Public Affairs, 1942.

Hartmann, H. Comments on the psychoanalytic theory of the ego. In *The psychoanalytic study of the child*, Vol. V. New York: International Univ. Press, 1950. Pp. 74–96.

Hartmann, H. Ego psychology and the problem of adaptation. In D. Rapaport (Ed.), *Organization and the pathology of thought.* New York: Columbia Univ. Press, 1951. Ch. 14.

Hyman, H. *Political socialization.* Glencoe, Ill.: The Free Press, 1959.

Inkeles, A., & Bauer, R. A. *The Soviet citizen.* Cambridge: Harvard Univ. Press, 1959.

Isaacs, H. R. *Scratches on our minds; American images of China and India.* New York: John Day Co., 1958.

Kardiner, A., *The psychological frontiers of society.* New York: Columbia Univ. Press, 1945.

Katz, D., McClintock, C., & Sarnoff, I. The measurement of ego defense as related to attitude change. *J. Pers.*, 1957, 25, 465–474.

Katz, D., Sarnoff, I., & McClintock, C. Ego defense and attitude change. *Human Relations*, 1956, 9, 27–45.

Kluckhohn, C. Politics, history, and psychology. *World Politics*, 1955, 8, 112–123.

Lane, R. E. *Political life: Why people get involved in politics.* Glencoe, Ill.: The Free Press, 1959.

Langer, W. L. The next assignment. *Amer. Historical Rev.*, 1958, 63, 283–304.

Lasswell, H. D. *Psychopathology and politics.* Chicago: Univ. of Chicago Press, 1930. (a)

Lasswell, H. D. *World politics and personal insecurity.* New York: Whittlesey House, 1930. (b)

Lasswell, H. D. *Power and personality.* New York: W. W. Norton, 1948.

Lazarsfeld, P. F., Berelson, B., & Gaudet, Hazel. *The peoples choice.* New York: Columbia Univ. Press, 1948.

Leites, N. *A study of Bolshevism.* Glencoe, Ill.: The Free Press, 1953.

Leites, N. Psycho-cultural hypotheses about political acts. *World Politics*, 1958, 1, 102–119.

Lerner, D. *The passing of traditional society.* Glencoe, Ill.: The Free Press, 1958.

Lipset, S. M. *Political man: The social bases of politics.* New York: Doubleday & Co., 1960.

Lynd, Helen M. *On shame and the search for identity.* New York: Harcourt, Brace & Co., 1958.

McClosky, H. Conservatism and personality. *Amer. Polit. Sci. Rev.*, 1958, 52, 27–45.

Mead, Margaret. The study of national character. In D. Lerner and H. Laswell (Eds.), *The policy sciences: Recent developments in scope and method.* Stanford, Calif.: Stanford Univ. Press, 1951.

Potter, D. *People of plenty.* Chicago: Univ. of Chicago Press, 1954.

Riesman, D. *The lonely crowd.* New Haven: Yale Univ. Press, 1950.

Sarnoff, I., & Katz, D. The motivational bases of attitude change. *J. abnorm. soc. Psychol.*, 1954, 49, 115–124.

Smith, M. B., Bruner, J., & White, R. W. *Opinions and personality.* New York: Wiley, 1956.

(Manuscript received December 6, 1960.)

The theory that what we find out about the real world is dependent upon the terms we use as well as upon the nature of reality is a major contention of this article. Since the words of a proposition are no more dispensable than its empirical content, what is necessary is not the elimination of definitions and their role, but the establishment of a satisfactory "balance" between what is conceptual and what is concrete. Here three such possible "balances" are illustrated in arguments about international relations.

CATEGORIES, CONCEPTS, AND REASONING IN THE STUDY OF INTERNATIONAL RELATIONS

by Richard N. Rosecrance

University of California, Los Angeles

> If the theorist starts with an *a priori*, or, what is more likely, an anachronistic method, he will only discover in history what he has himself unconsciously put into it. He will be able to prove to his own satisfaction, doubtless, what he set out to prove, but his material will not be history, and his conclusions will not be social science (COBBAN, 1956, p. 17).

IN the epigraph to this paper, Alfred Cobban, the distinguished English social theorist, marks an error which is familiar to all of us. Sometimes we intend to make empirical assertions, but end only in making logical ones; we ourselves "put into history" what we seek to discover there. Other times we want to show that an opponent is wrong, but compose an argument which proves him to be self-contradictory. If we show the latter, we cannot have proved the former, for logical processes do not give grounds for empirical ones. In these cases we have mistaken "analytic" for "synthetic" assertions. Instead of empirically discovering that "John is a brother," we have contented ourselves with the truism "A brother is a male sibling."

But in international relations, as in other social disciplines, the identification of synthetic and analytic arguments is not the only source of possible misapprehension. Sometimes our arguments are indubitably empirical (synthetic), and yet the conclusions we draw depend inordinately upon definitions of central terms. This is particularly the case where terms are (1) defined in a very general but precise fashion, or (2) loosely defined and capable of extension in usage. If, for example, I assert that all human motivations are either egoistic or altruistic and define as altruistic an action which is performed despite a conviction that it conflicts with one's desires, feelings, and even moral sensibilities, it would be no surprise to find that most actions are egoistic in inspiration. In such a case the true statement "X (action) was motivated by egoism" would be a synthetic not an analytic proposition, but its truth would depend inordinately upon the very broad definition of "egoism."

Broad or extensible definitions also directly influence the empirical measures we take to prove assertions. Suppose two definitions of a single term K are used in investigations in the international field, and suppose also that one of those definitions is both precise and narrow and that the other is not. To establish the presence of K on definition (1), characteristics (a), (b), and (c) must be found to exist; to establish the presence of K on definition (2), only characteristic (a) must be ascertained. In order to make true statements about $K(1)$, therefore, two additional operations must be

performed on reality than would be necessary to make true statements about $K(2)$. In this manner, general or Protean definitions tend to influence the empirical steps we take to establish propositions; such definitions limit empirical operations compared to those which are both narrow and precise.

This is not to say that narrow definitions have no impact on the conclusions we draw; no definition ever conceived is so pure or elevated as to exert no influence upon our conclusions. What we find out about the empirical world depends upon the definitions of the terms we use and upon the nature of empirical reality, and the first category is no more expendable than the second. What is necessary is not the elimination of definitions and their role, but the establishment of a satisfactory balance between term and reality, between what is conceptual and what is concrete. And we must focus attention not only on the interaction of the two strata, but on the nature of the balance achieved. On the one hand, we may use specific definitions that leave the great burden of our conclusions to be provided by the empirical realm as delineated by empirical research; on the other hand, we may use general definitions that themselves provide the great burden of the conclusions we reach. That both types of definition exist can scarcely be doubted. We often use concepts which cannot be applied to experience without considerable empirical investigation. We often use concepts whose application to reality requires no practical research or study. Some modern philosophers have believed that traditional philosophical puzzles concerning free will and determinism, mind and body, the problem of sense data, and so on, are of the latter type; that is, they revolve around the meaning of central terms, not around the experiential world (Wittgenstein, 1922; Cohen & Nagel, 1934). And in political science itself definitions may be divided into two broad classes *with gradations in between*: some of our definitions are contingent, in that their application can only be determined by pragmatic inquiry; some are necessary, in that their application is determined in advance by logic.

What I shall be concerned with in this article is the proper "balance" to be attained between conceptual and empirical factors, given certain intellectual objectives. Specifically, I should like to illustrate the respective uses of concept and reality in three kinds of statements about international affairs and totalitarian dictatorships and to comment briefly upon the balance attained in each instance. The first has to do with the use of different definitions of a single term; the second concerns the use of terms which have an especially influential or partially determinative impact upon the result achieved; the third refers to that usage of terms which causes a transition from empirical (synthetic) to logical (analytic) statements.

I

It is not always recognized that although each scientific investigator has unrestricted latitude in defining his terms, he must be especially careful if general usage has already established meanings for them. For if his terms are in general use, he must distinguish his definitions from those employed in the public at large. This admonition raises special problems for the social scientist because he rarely has an opportunity to devise terms which are not in common parlance. Occasionally, of course, an investigator may talk of "pluralistic or amalgamated security communities" with relative impunity; by and large it remains true, however, that the social scientist must muddle through with a terminology not of his own creation while the physical scientist enjoys considerable scope in building his central notions. A physical scientist can encounter difficulty only when he transposes meanings within his own unique terminology. The social scientist may be misled not only by his special terms; he may not clearly differentiate his own usages from the conventional ones of ordinary language, and he may not always perceive the differences in various conventional usages. Definitional regularity is more difficult to attain in social than in physical science.

The result of the use of variant meanings

of a single term is often a shifting balance between what is conceptual and what is concrete. While one definition of a key term may be quite narrow, another may be correspondingly broad; in the one case, the truth of a statement embodying the term may depend upon extensive empirical inquiry; in the other, it may largely be assured in advance by the scope and reference of the central concept. "Power" is a case in point, for it is a peculiarly Protean word. One can think immediately of at least three conventional senses of "power" (and these certainly do not exhaust its meanings). There is firstly the sense in which "power" is used in the phrase "power politics." Here "power" denotes bullying, conflict, the threat of force, or force itself. There is a second, more general sense in which "power" suggests competitive activity or rivalry. Thirdly, there is an even more general sense in which "power" is synonymous with "capability." This sense is dispositional, while the others depict certain kinds of activity. These three uses of "power" may be exemplified in prevalent arguments about international relations and domestic affairs.

One of the exemplary purposes of the "power analysis" of international affairs was to indicate the limitations of the policies of idealism and appeasement pursued in the period between World Wars I and II. It was necessary to make clear to the general public that international politics was indeed "power politics" and that wishful thinking and a boundless willingness to compromise would not transform the international environment (see Morgenthau, 1956, p. 6). One eminent theorist linked power drives among nations to an intrinsic egoism in human nature which impelled one individual to assert himself at the expense of another (Carr, 1946). The result of egoistic assertion in international society was conflict, aggression, and sometimes war.

Essentially the same use of "power" appears in historical arguments about the conditions of international relations. It has been said, for instance, that after the First World War, universal moral principles failed to restrain competition among nations. Before 1914, nations sometimes acted in accord with universal moral precepts; with the Great War, however, nations began to clothe their national interests in parochial morality and a period of "nationalistic universalism" arrived (Morgenthau, 1956, pp. 220–234). If the gentlemen-diplomats of the eighteenth century and Prince Bismarck in the nineteenth had heeded a code of morality in their dealings with other states, "nationalistic universalism" represented the final triumph of power-seeking over moral restraints. Hence, international conflict reached a new fury; a period of unequivocal "power politics" was ushered in.

A second sense of "power" comes to us from the analysis of domestic society. We have often heard that all politics manifest a struggle for power: the mother-in-law conflict, the struggle for pre-eminence in professional organizations, the rivalry expressed in competitive examinations, the competition between industrial and labor organizations, the drive for wealth and prestige, and the political raillery of periodic elections all evince the struggle for power in domestic life (Morgenthau, 1956, Ch. 3). Here "power" denotes "competition," the limited conflict of civil society as contrasted with the relatively unlimited conflict of international relations. If the logical extension of the first sense of "power" is war or national antagonism, the second sense indicates the sheathed rivalry of civil competition.

A third use of "power" is different from the first two. It has been said that: "Whatever the ultimate aim of international politics, power is always the immediate aim. Statesmen and peoples may ultimately seek freedom, security, prosperity or power itself. . . . But whenever they strive to realize their goal by means of international politics, they do so by striving for power" (Morgenthau, 1956, p. 25). Although certain complications are introduced in this argument by the phrase "by means of international politics" its burden seems clear. It would seem to suggest that if one wants to do something, he must first seek the power to do it, and this usage introduces a new sense of the term. If one seeks ends, one must seek the "capability" (or power) to achieve them. The pursuit of power then becomes char-

acteristic of social behavior that is end-oriented. On this definition, "power" does not refer to an activity, but to a dispositional quality; it is potential more than actual.

When each of the three senses of "power" is employed in the assertion "international politics is a struggle for power,"[1] the balance between the definitional and the real shifts from one to another. The first sense of "power" is the most narrow and specific; if not only denotes competition, but the most virulent form of competition—strife, force, or war. Certain elements of antagonism must be present in an international situation for it to be designated as "power politics." Whether those features exist or not cannot be determined in advance by the central concept; it can be shown only by investigation into the present or past conditions of international relations.

The second sense of "power" is less specific. "Competition" is a much broader term than "hostility" or "antagonism," and is therefore characteristic of a wider range of social experience; the criteria for the existence of "competition" are vaguer than they are for "hostility." When this meaning of "power" is inserted into the proposition "international politics is a struggle for power," the balance between concept and empirics is altered. "International politics is conflict, strife, and bullying" leaves a considerable latitude to the empirical realm because the concepts are themselves restricted; "international politics is competition" narrows that latitude because "competition" has an intrinsically broader reference than "conflict" or "strife." While "conflict" and "strife" are forms of "competition," the reverse is not the case.

The third sense of power is even more general. "Power" as "capability" has an even more inclusive meaning than it does in its other senses. That "international politics is the pursuit of capability" is nearly plausible on definitional grounds alone. Because "capability" is a very general concept, the phenomena which it embraces are many and varied; the need for precise empirical tools to show that "international politics is the pursuit of capability" is thus less compelling, and the assertion can be verified with minimal recourse to empirical investigation. "Power" in this third sense, then, is broader than in the other two, and because it is broader, its application is as much (or more) the result of its intrinsic extent as a concept as it is of the character of the real world.

The shifting balance of concept and reality according to the meaning used has an important impact upon the truth or falsity of assertions. Whether "international politics is a struggle for power" or not will depend upon the meaning of "power" employed. It is possible that valid propositions formulated on the basis of definition (1) of term *X* may be false when they are based upon definition (2) of term *X*. It certainly may be that "international politics is conflict, strife, and bullying," "international politics is competition," and "international politics is the pursuit of capability" are all true assertions; on the other hand, they may not all be true. In any case it can scarcely be denied that the meaning of the statement "international politics is a struggle for power" varies significantly with the definition of "power" that is used.

II

A second usage of terms interesting to the social scientist is concerned with definitions which have a peculiarly influential or partially determinative impact upon the conclusion drawn. The consequences which flow from the employment of such concepts may be indicated in several examples. One student of modern totalitarianism writes: "The consolidation of Stalin's absolute power, the ruthless purge of party dissidents, the strengthening of the bureaucratic and hierarchical features of the regime, the expansion in the role of the secret police, and the rise of forced labor, advanced hand in hand with forced-draft industrialization and militarization" (Fainsod, 1953, p. 109). And again: "The formula of totalitarian rule as it took shape under Stalin's ministrations was a complex one. It represented, in one aspect, a drive to safeguard his own security

[1] All three senses of "power" are frequently employed in this formulation.

by obliterating all actual or potential competing centers of power" (Fainsod, 1953, p. 109). These passages suggest the author's belief that one important means of seeking power in a totalitarian society is through the multiplication of restraints upon the civil populace; perhaps more important, it suggests that where evidence is found of the imposition of restraints, one may conclude that the imposer is pursuing a "power" policy. But if in the instance above Stalin sought to achieve power by regimentation and compulsion, he also allowed scope to the forces of Russian tradition. Fainsod writes:

> . . . Stalin also endeavored to identify totalitarian rule with the forces of tradition and respectability in Russian life. This search for stability and legitimacy took him along strange and devious paths. It led to a drastic reorganization of the educational system, the abandonment of its early experimental and progressive features, and its transformation into an authoritarian instrument to instill devotion to the regime and to prepare youth for their appointed roles. . . . It manifested itself in a restoration of the authority of the family, in restrictions on abortions, in encouragement of child-bearing, and in tightening marriage bonds. It produced an uneasy *de facto* "concordat" with the church in which the political loyalty of the clergy and their communicants was exchanged for a precarious toleration of religious practices. It expressed itself in a striking rehabilitation of patriotism as the cohesive force of Soviet society and sought increasingly to present the Soviet regime as the legitimate heir of the best in the Russian past (Fainsod, 1953, pp. 109–110).

This passage suggests that Stalin did not rely alone on force and restriction to achieve and consolidate his power; if he increased his control by weeding out competing centers of power, he also strengthened his hand by conceding greater influence to traditional Russian social institutions, the family, the school, and the church. That both the policy of allotting leeway to Russian tradition and that of enforced restriction illustrate the pursuit of power by Soviet totalitarians is indicated as follows:

> In the tangled web of interacting influences which contributed to the "traditionalist" restoration that began in the thirties, probably the most basic was the leadership's drive to stabilize its position and to extend its power in both the Soviet and world arenas. Rapid industrialization and the revival of traditional nationalism were both directed toward this larger purpose. Most of the significant "conservative" innovations which accompanied them were introduced because of the same dominating preoccupation. . . . The emergence of Stalinism in the conservative garb of the party of tradition and order was once described as a "Great Retreat." It gave this surface appearance, yet the phrase was delusive. In utilizing such traditional pillars of authority as nationalism, family, and the church, Stalin did not surrender power. He consolidated it (Fainsod, 1953, pp. 115–116).

The question raised by this analysis is not whether in carrying out his "power" policy Stalin did certain things; it is whether the things that he did give us warrant to claim that he followed a "power" policy. Did his actions in the Soviet industrial sphere, and in the spheres of family life, traditional nationalism, and church affairs constitute a "power" policy or did they exemplify other motivations? There is little doubt that the theorist concludes that "power" is the dominant aim. Summarizing his findings, Fainsod (1953, p. 87) says: "To understand the turnings and twistings of Soviet policy, the advances and the retreats, the continuities and the reversals, it is essential to comprehend the preoccupation with power which has motivated the responsible Bolshevik leadership *at every crucial stage* [my italics] in the development and consolidation of the Soviet political system." The definition of "power" suggested in this analysis is a broad one; if the policy of the carrot and the policy of the stick manifest in equal degree the pursuit of power, intermediate policies of quasi-conciliation or quasi-repression must also evince power considerations. The only occurrences which would not fall under the power category are those which involve the pursuit of ideological goals, and this suggests that "power-ideology" is viewed as an exhaustive categorization (Fainsod, 1953, Ch. 4). Of the two concepts, "power" is the more broadly defined.

It may be maintained that the relatively extended meaning of "power" has an important and influential impact upon the conclusion of the argument; because of the breadth of "power" and the slenderness of "ideology," one would "expect" power to be

the major goal of totalitarians, and this leads one to conclude that the balance between concept and reality has been weighted in favor of the former. The same basic features exist in the analysis of international politics. Whether most nations seek power or certain "moral" ends depends in part on the way the respective terms are defined. Those theorists who define "power" broadly and "morality" narrowly will usually be led to the conclusion that "power" is the overriding goal in international politics. If the theorist begins by defining "power" narrowly, and juxtaposes to this narrow definition broad constructions of "ideology" or "morality," it is likely that the vast burden of events in the study of totalitarian dictatorships or international politics will fall into the "ideology" or "morality" category. In either case, if the dichotomy "power-morality" or "power-ideology" were not regarded as exhaustive, quite different results would follow.

If the definition of "power" exerts an influential or partially determinative impact upon the degree of "power-seeking" in modern totalitarianism or international affairs, the concept of "equilibrium" has a similar effect. The definition of "equilibrium," embracing "general and partial, unique and multiple, stable, neutral and unstable, long-term and short-term, perfect and imperfect" (Liska, 1957, p. 12) phases, is quite general and its very breadth ensures its application to a wide range of international phenomena. As its proponent states: "One fact emerges clearly from the analysis: it is the persistent reassertion of the balance or equilibrium or power as the irreducible if changing central dynamics of international relations" (Liska, 1957, p. 193). Whether or not the central term of analysis can be applied meaningfully to a large number of situations depends as much upon its definition in the first place as it does upon international reality. The balance of the conceptual and the concrete is again in favor of the former.

A very different but equally interesting usage of terms occurs in the explanation of integrative processes in international politics. According to two recent studies a "pluralistic security community" is found to exist where there is the "attainment within a territory of a 'sense of community' and of institutions and practices strong enough and widespread enough to assure for a 'long' time dependable expectations of 'peaceful change' among its population" (Deutsch et al., 1957, p. 5) and where that territory embraces two or more separate and independent governments; expectations are "dependable" when they are "geared not only to the current load of mutual transactions and potential conflicts among participants, but . . . to any increase in the volume of such transactions. . . " (Deutsch, 1954, p. 40). "The combined territory of the United States and Canada is an example of the pluralistic type. Its two separate governmental units form a security community without being merged" (Deutsch et al., 1957, p. 6). "Thailand and Bolivia," on the other hand, "have no significant conflicts, but neither have they any significant mutual transactions. They do *not* constitute, therefore, a pluralistic security community."[2] The amalgamated community, in contrast, consists in "the formal merger of two or more previously independent units into a single larger unit, with some type of common government after amalgamation. This common government may be unitary or federal. The United States today is an example of the amalgamated type" (Deutsch et al., 1957, p. 6).

The definition of "pluralistic security community" rendered above is of a different kind from the ones previously considered. It is an extremely circumscribed definition, and would apply not to the majority, but to a small minority of states of the world. "The cases of pluralistic security communities are. . . neither numerous by their definition nor is it obvious or undisputed that they should be numerous in fact." In addition, in contrast to other definitions considered, there are "empirical criteria by which relatively small number of pluralistic security communities are to be identified".[3] The respective balance of concept and

[2] Professor Karl W. Deutsch's comments upon a previous draft of this paper.

[3] Professor Karl W. Deutsch's comments.

reality in determining the *existence* of pluralistic security communities is weighted heavily in favor of reality; no large concepts here make the existence of such communties the result of definitional postulation. It may still be maintained, however, that the definition of "pluralistic security community" has a vital impact upon the conclusions derived about the respective *success* of pluralistic and amalgamated communities in eschewing war among participating units. The criterion of a pluralistic security community is the existence of "dependable expectations of 'peaceful change' among its population" and "a high level of mutual transactions." The criterion of an amalgamated community, on the other hand, is the existence of a common political bond and central decision-making center. It is relatively easy to imagine cases where territories under a single political bond could find occasion for war with one another; it is more difficult to conceive of situations in which territories harboring not only mutual expectations of peaceful change but a high level of reciprocal transactions would come to blows. This is *not* to say that "pluralistic security communities" cannot break down in strife; even where expectations of peaceful change and a wide range of mutual transactions exist, there may still be war. Such a breakdown, however, would be the unexpected, not the normal case, while the dissolution of the amalgamated bond is today as much the rule as the exception in international politics.[4] Indeed, the force of nationalism in present-day politics militates as often against amalgamation as it does in favor of it. It is not unnatural, then, that "pluralistic security communities" are "easier to preserve than their amalgamated counterparts," and that there was "only one case of a pluralistic security community which failed in the sense that it was followed by actual warfare between the participants," in contrast to the "seven cases of amalgamated security communities that failed" (Deutsch et al., 1959, pp. 29–30.) Significantly, one only applies the designation "pluralistic security community" when the prospect of war among potential constituent units is already remote.

Relative to the *existence* of such communities, the definition of "pluralistic security community" is very narrow, and the balance weighted in favor of reality; relative to the *success* of such communities in avoiding intramural strife, the definition is not broad but it is nonetheless weighted in favor of the conceptual realm. It may be suggested, then, that the conclusion that pluralistic security communities are more successful in avoiding wars among participating units than amalgamated ones flows as much from the respective definitions employed as from the nature of the empirical world.

III

A third usage of terms interesting to the social scientist occurs in the transition from empirical to logical statements. This happens when concepts come to replace the empirical realm. Many of the arguments commonly heard about international relations are of this type. We have often held, for instance, that:

(1) All nations seek to do things in the international world.
(2) In order to do something, it is necessary to have the power to do it.
(3) Therefore, all nations must seek power.[5]

There are two possible interpretations of this argument: the first is that its purpose is modular or heuristic; the second is that it is intended as an introduction to an empirical account of what nations do. According to the first interpretation, what is being said is that to seek *any* end is *ipso facto* to seek power. In our language it is impossible to conceive of doing anything without having the power to do it, for the idea of "action" involves that of "power." If the notion of

[4] In rebuttal to this point, it has been argued that what is "expected" or "normal" is a function of the individual investigator; while some may "expect" pluralistic security communities to maintain themselves, others may "expect" them to break down. To be sure. But a single investigator cannot both "expect" pluralistic security communities to maintain themselves and announce their maintenance as a striking or surprising *dénouement*.

[5] This argument appears in Morgenthau (1956, p. 25).

"doing anything" involves definitionally or logically "having the power to do it," the idea of "seeking to do anything" must similarly involve "seeking the power to do it." The result is a logico-deductive analysis which must hold by the very nature of the terms used; it is the elaboration of an ideal type or "model."

The second interpretation sees the argument as the genesis of an empirical theory. It sees an analogy between what states do in attempting to conserve or increase their power and what an individual does in striving to build up his generalized "purchasing power" as an alternative to expending a part of it on the purchase of consumer items; it is similar to the refusal of people to invest when the interest rate is low because of high "liquidity preferences."[6] On this interpretation nations always face a choice between the alternatives of using or conserving their power; in either case, power is employed to effectuate ends, but at any juncture a state must choose between actually utilizing it (and thereby diminishing it) or consolidating it. On this interpretation the argument invites us to inquire into the situations in which the second choice will be made in preference to the first. The difference between these interpretations is radical; on the one hand, empirical investigation is obviated, on the other it is enjoined.

Secondly, it is sometimes suggested that although types of international activities have no reference to power considerations (for example, certain humanitarian, protocol, and economic activities), they should be excluded from the sphere of international politics. (See Morgenthau, 1956, pp. 25–26.) On this argument, politics *is defined* in terms of the struggle for power; where "power" is not present, "politics" is not involved. Where international activities do not or imperfectly manifest power considerations, international politics is not implied. There is little doubt that this is a deductive argument, depending solely upon the definition of terms, and its result is an "ideal type" analysis.

Thirdly, it is often suggested that since social affairs exemplify a struggle for power, international affairs, a part of social affairs, must also evince a power struggle. (See Morgenthau, 1956, p. 31.) This argument is less obviously an ideal typology. The claim that international affairs are a struggle for power rests upon an assertion that social affairs, generally, are a struggle for power, and this leads us to an examination of the latter contention. Such a claim may be defended against examples of the most self-sacrificing behavior, for self-sacrifice may ultimately be a superior means to power justified by the maxim "The meek shall inherit the earth." Altruistic acts, however disinterested, family life, however loving, may be conceived in terms of power as a motivating objective. Even martyrdom can easily be explained in terms of the pursuit of power in human affairs. Indeed, it eventually becomes evident that this logical extension of Adlerian psychology can admit of no exceptions; one cannot even conceive of hypothetical cases which would refute the contentions of the power theory. Consequently, the statement "all social affairs are a struggle for power" is *unfalsifiable*, because it is impossible to imagine one theoretical instance in which it would be false.[7] But the very nature of unfalsifiable statements is logical or deductive, not empirical. An unfalsifiable statement is formulated in such a way that no empirical fact can affect is validity, and if no empirical fact can affect is validity, it must be founded not on empirical but on logical principles. Once it becomes clear that the assertion "all social affairs are a struggle for power" is logical (analytic), it is clear that the derivative premise "international affairs are a struggle for power" is also logical (analytic). It may be suggested, then, that the third argument is a "model" or "ideal type" and is thus similar in several respects to the first and second arguments. In all three cases interpretations lead us to believe that logical propositions are being enunciated.

[6] I owe this point to Professor Karl W. Deutsch.

[7] Professor K. R. Popper says, for example, "So it is, I hold, the possibility of overthrowing it, or its falsifiability, that constitutes the possibility of testing it, and therefore the scientific character of a theory. . ." (1950, p. 444).

IV

Various arguments about international relations, totalitarian dictatorships, and social relations have been reviewed in this essay. The questions raised by these arguments are complex and cannot be unraveled easily by a single investigator. But perhaps some major features of these issues can be clarified in the hope that further inquiry will produce more satisfactory conclusions. It is, I believe, inescapable that words play a role in the results we obtain; how broadly or how narrowly we define a concept has an immediate impact upon the conclusions we derive when the term is applied to experience. What we find out about the world depends upon the kind of definitions we use as well as upon the nature of reality. If it is accepted that there is always a balance (in one direction or another) between term and reality, the most significant issue becomes the kind of balance to be achieved. Are narrow definitions, leaving the greater scope to "reality," preferable to broad usage of terms or is the reverse the case? The answer, I believe, depends in large measure on the purpose of analysis.

For years the study of international relations sought unifying concepts, concepts which could be applied in the entire range of international experience and which could order a mass of otherwise unco-ordinated phenomena. If we were to learn something about international relations we had to have these concepts or at least quite a number of them; "power," balance of power," and "equilibrium" have been able to engross a very large number of international occurrences, and a token of this development of unified "theory" on the basis of general concepts has been the wide acceptance of the "power" approach to international relations. "Power," it is fair to say, has been to international relations what "economics" (in the Marxian sense) was to society and what "Eros" (in the Freudian sense) was to psychology. If our intention is, then, to formulate unified and univocal theories of international society after the manner of Marx and Freud, our usage must be broad. If our intention is to develop ideal typologies, our definitions must not only be broad, they must be logically preemptive; they must be universal. As the economists must assume free competition in order to discuss models of industrial order where an infinitely elastic demand curve exists for each firm, so it may be necessary to assume the universal quest for power in order to discuss in modular form the conditions of alliance and opposition among sovereign states.

If, on the other hand, our purpose is not further comprehensiveness in the study of international relations, nor the adumbration of definitions which will facilitate "type" analysis, narrow definitions may better serve our purposes. If definitions can aid our understanding of the real world, they can also pose a mighty hindrance, and if our purpose is not to bring an everlasting order out of chaos but to delineate the conditions of chaos, precise conceptual tools will be helpful. It may be said that what we find out about the empirical world depends upon the "operations" we are forced to make upon it, and in this surgery a probe may be more useful than a cleaver. Broad definitional formulations reduce the number of operations we must go through in order to prove our case; narrow ones actually require a greater number of operations. If "operationalization" then is desirable, narrow, operational definitions may facilitate our understanding of the international world. (See Bridgman, 1936, ch. 2.) One should not expect new comprehensive formulations to emerge from this relatively detailed investigation; indeed, it must erode comprehensiveness and order at least in its initial stages, but the result will permit a more careful attention to the vagaries and variety of reality than is possible with the current conceptual categories.

In terms of this kind of empirical research, the interchangeable use of different definitions of a central notion is a roadblock, not a clearly marked highway. Protean terms make investigation difficult; we must always be sure what sense we are using, and we must make sure that arguments which we formulate in terms of a single use of a term are not extended to a second use without careful reinvestigation. Where there is a shifting

balance of concept and reality, the invitation to inconsistency is difficult to resist. The development of terms which are general or which exert a peculiarly influential impact upon the results we derive advances the purposes of comprehensiveness, but it circumscribes conclusions of detail. Universal definitions are prerequisite to hypothetical analysis, but they are an absolute impediment to empirical research.

Most of all, perhaps, the danger we must keep before us is the confusion of analytic purposes, and the adoption of methods and concepts appropriate to one sort of investigation for use in another. We cannot begin with the announced purpose of proving empirically that international politics is a struggle for power and then develop logical or deductive arguments to that end; we cannot avow a desire to probe delicately beneath the surface of international reality and then employ general definitional concepts for that purpose; we cannot insist upon a single usage of a crucial term and then employ other uses. In more general terms, the study of methods suitable for given research needs is perhaps more advanced in other disciplines than it is in international relations; let us hope that new inquiries will soon commence to remedy the defect.[8]

[8] The present author's forthcoming *Empirical Systems of International Relations* seeks to make a modest contribution in this direction.

REFERENCES

Bridgman, P. W. *The nature of physical theory.* Princeton, N. J.: Princeton Univ. Press, 1936.

Carr, E. H. *The twenty years' crises* (1919-1939). (2nd ed.) New York: St. Martins Press, 1946.

Cobban, A. The vocabulary of social history. *Polit. Sci. Quart.*, 1956, 71, 1–17.

Cohen, M. R., & Nagel, E. *An introduction to logic and the scientific method.* New York: Harcourt, 1934.

Deutsch, K. W., *Political community at the international level.* Garden City, N. Y.: Random House, 1954.

Deutsch, K. W., Burrell, S. A., Kann, R. A., Lee, M. Jr., Lichterman, M., Lindgren, R. E., Loewenheim, F. L., & Van Wagenen, R. W. *Political community and the north atlantic area.* Princeton, N. J.: Princeton Univ. Press, 1957.

Fainsod, M. *How Russia is ruled.* Cambridge: Harvard Univ. Press, 1953.

Liska, G. *International equilibrium.* Cambridge: Harvard Univ. Press, 1957.

Morgenthau, H. J. *Politics among nations.* (2nd ed.) New York: Knopf, 1956.

Popper, K. R. *The open society and its enemies.* Princeton, N. J.: Princeton Univ. Press, 1950.

Wittgenstein, L. *Tractatus logico-philosophicus.* London: K. Paul, Trench, Trubner & Co., 1922.

(Manuscript received January 4, 1961.)

The 'well-rounded man' has become the organization man, the fraternity brother, or the man who is so well-rounded that he rolls wherever he is pushed.

ALAN SIMPSON, "Liberal Education in a University"
Washington University Magazine

The study of various aspects of organization theory has been undertaken by this author in previous articles in *Behavioral Science*. In the January, 1958 issue, Models for Man and The Organization of Space were discussed, while in the July, 1959 issue, problems relating to Decision Making, Planning, and The Steady State were explored. Here the author emphasizes synthesis of organization, as opposed to the earlier descriptive analytical studies, with a review of recent literature on simulation studies of social organizations. Simulations thus far have tended to concentrate on such things as management games and war games—a narrow range considering the ultimate possibilities of their use in understanding behavior. What educational and training purposes can simulations be used for? How can we further expand this new method of studying social organizations?

EXPLORATIONS IN THE REALM OF ORGANIZATION THEORY. IV: THE SIMULATION OF SOCIAL ORGANIZATION

by Richard L. Meier

Mental Health Research Institute, The University of Michigan

THE previous explorations in this series (Meier, 1958a, 1958b, 1959) concentrated upon a few rather successful and surprising items. Each was singled out for attention primarily to clarify other items and simultaneously lay the groundwork for brief speculations about the next steps that might be taken toward the elaboration of a satisfactory theory of organization. Most of the studies reviewed before were descriptive and analytical—the product of many nonparticipant observers summarized and restated by an expositor–observer. The work that has to do with the *synthesis* of organization, as contrasted to analysis, description, and evolution, has been passed over without much comment thus far, mainly because the published work has been superficial, or extremely specific, or poorly integrated.

Perhaps the unsatisfactory nature of the work on synthesis of organization is due to the fact that this is a rather new field. In the more extensively developed fields, such as chemistry or physics, an ultimate test of the knowledge regarding a proposed structure is to demonstrate that it can be synthesized in an orderly manner from the basic elements. Thus the interest in synthesizing compounds like quinine, chlorophyll, or cyclooctatetraene among the chemists, and californium and positronium among the physicists, represents an elegant, almost ultimate, stage in explanation. Biologists similarly have been able to recombine genetic strains until they have achieved the wild type of an animal or plant species that had been domesticated centuries or millennia ago and the original wild type had become extinct. In a few instances the structures that had been deduced from analytic procedures have been shown to be faulty when synthesis had been achieved, but in recent times this has been quite rare. Thus the synthesis step, which usually requires the most advanced experimental techniques, is used to assure a community of investigators that their foundations are secure. The achievement of successful synthesis is proof of closure; it represents the end of a chapter.

Not all syntheses are as complex as the examples that have been given. Wöhler's synthesis of urea, a simple crystalline compound then obtained solely from biological sources, from inorganic raw materials, undercut the vitalist philosophy and, when backed up by other such indisputable syntheses, initiated a burst of investigation in organic chemistry.

Simulation is actually a kind of synthesis. It starts from simple elements and requires the construction of a system which closely approximates what has been observed in the real world. It should generate an illusion which people will immediately associate and compare with the real experience. A perfect simulation is one that cannot be distinguished from the previously established experience when using a series of "fair" tests. Thus in chemistry or biology it is possible for a scientific detective, using all of the clues at his disposal, particularly to trace impurities or strains, to distinguish the natural product from the synthetic. However, these methods are only conceivable *after* one knows that synthesis is possible and the routes that may be taken toward accomplishing it and thus cannot be considered as fair. Synthetic industrial diamonds five years ago, prior to the publication of the approach undertaken by General Electric scientists, would have passed for diamonds without question. Diamond merchants might have suspected a new and unannounced mine, but that is all, since the standard hardness and composition tests for diamond are met by the synthetic. At present, of course, they are readily distinguished from the natural by such properties as color, size distribution, shape, and other superficial properties.

Even ordinary, reproducible synthesis requires the joining of science with art. The art demanded is that of standard professional competence, not of the virtuoso or genius. The science, of course, is a systematic set of concepts shared by a body of investigators. As we know, it doesn't pay to research everything within the ken of man, but only those features which are crucial, or puzzling, or matters of mutual interest. Thus whenever a scientist or a promotor in the social arena proceeds to synthesize an element of organization, the handbooks and past publications provide only a part of the necessary information. The remainder must be introduced through artful improvisation.

Recently the studies on the synthesis of social organization have shown signs of progressing into a new stage of sophistication. Remarkable capabilities for growth and redirection in the social sciences are being created. Technique is improving and the questions that are raised in the course of some investigations refocus attention upon descriptive and analytical studies. Thus the term "organization" in a title attracts attention and will add to sales; it is being appropriated for a more diverse range of subject matters than ever before. The glamor is coming to be shared to some extent by old terms such as "administrative theory," "control," "systems," and the like if these are accompanied by a few clues which promise that the contents are not entirely a rehash of the classics.

The output of books and monographs on subjects related to organization is increasing very rapidly. My private estimate is that it is doubling every three years. We are already very close to the point where no single person can keep up with the body of published work and still spend some time making contributions. We must find various means for being selective. Therefore, it seems worthwhile at this time to undertake an elimination contest, one that subjects the newest contributions to a set of fairly rigorous tests that are suggested by the needs of the synthesizers. The failures can be as rewarding at this stage as the successes; they too give a sense of direction, a basis for choice, and an opportunity to economize upon effort.

What are the tests we shall apply to the current output? The first rule is that there should be concern with some principles, elements, or techniques involved in the modeling of organization and that relationships are clarified. Each case study, for example, should bring to the fore all the essential variables, and establish a basis for estimating parameters, so that it becomes possible to explore, with simulation, the implications of actions *not* really taken at the time the case was observed. When all major factors have been identified and incorporated into a simulation system, the history of the original system can be reproduced on a small scale, enabling us to understand better the role of chance factors, the consequences of errors that could have been induced by the stresses of the situation, and the utility of various

clues for diagnosing similar situations when they are encountered at a later date. As we shall see, very few case studies are thoroughgoing enough, or have sufficient breadth, to permit them to be used for creating a simulation. On the other hand, a useful theoretical discussion might be rather specialized; it could pinpoint one or more crucial regions for the empirical testing of "working models." Normally it would supply criteria which would determine how close the model is to the real thing, or vice-versa. Thus our search of the current literature aims to discover complete descriptions and ideas which have a reasonable likelihood of large payoff in the process of synthesizing social organization, realizing that simulation is not a once-for-all-time achievement but a process that goes through successive stages. We are also, of course, interested in the uses to which simulations are put, and the inferences that are drawn from their behavior.

Admittedly, this is a cavalier treatment for books that were planned for the most part some time between 1952 and 1956, when the concept of simulated social organization was no more than a gleam in the eye of a handful of investigators, and even the idea of controlled experiments was still quite novel. Nevertheless, readers of the 1960's have new interests, and the synthesis of organization, either on a laboratory scale or in the real world, is one of these. Thus the contributions will not be judged in the terms that their authors, editors, and publishers intended, but according to values being developed by one branch of successors.

In taking this position, we must always be conscious of one additional challenge that faces the new generation of social investigators. It stems in large part from recent advances in computing ability. If the capacity of both the instruments and the knowledge of programming technique are taken into account, it appears that information-processing capability has increased by a factor of ten times at most research centers in the past five years and promises to do so again in the next five. The new capacity can be used for grandiose analytical schemes, such as the inversion of input-output matrices for a national economy, or a gigantic factor analysis, but it seems likely that we are approaching a period of diminishing returns in that direction. The data so processed are seldom good enough to be worth interpreting that finely. The more imaginative uses of this hundredfold increase in the ability to process information involve the manipulation of interactions between elements of a prototype organization, discovering which sequences and aggregates of interactions are more life-like or otherwise interesting. The real contribution of the added computing capacity is that we are enabled to move from static structures to dynamic patterns.

The path through the literature is full of many twists and turns. I have taken the liberty of reviewing books and monographs under two or more subheadings when that seemed worthwhile, thus permitting some sense of ordering and direction. The evaluation begins with the collection of facts and the changes in the methodology of fact-gathering introduced by the possibilities of simulation. What was competent science in the recent past often turns out to be inadequate now. From there we move on to the system-defining task, concerned with the first steps in understanding the dynamics of an isolable unit of organization, and with experiences prior to, but sometimes including, an attempted simulation of the system. These gropings are still so inconclusive that one is impelled to look for guidance from recent volumes on the theory of organization. Finally, it is recognized that this review of current effort reveals a surprising one-sidedness in the early attempts at synthesis. Fairly explicit proposals can now be made for extending techniques of simulation to other very promising areas of social organization and simplifying them so that they can be introduced rather early in the educational process.

GETTING THE FACTS

The emphasis upon "broad, balanced, thoroughgoing case studies of organization" as a prerequisite for simulation does not hint at the complications that follow. Human organization is so complex that elaborate strategies for selecting and interpreting facts and events are required in order to achieve

an economic representation. A biologist can imagine himself a bacterium making observations upon the surrounding tissue and infer thereby the problems of scale faced by the social investigator when characterizing an organized environment. It seems appropriate to begin this survey, therefore, with a new collection of articles entitled *Human Organization Research* (Adams & Preiss, 1960) on the subject from a journal explicitly devoted to methodologies of field research. *Human Organization* is an organ of the Society for Applied Anthropology.

Despite the care taken in assembling this smorgasbord, and despite also an interesting attempt at updating the pieces by offering the authors a chance to extend or revise their prize contribution, the total effect is disappointing. The offerings are not balanced or substantial enough for the heavy work at hand. At least a half-dozen of the 32 pieces reflect the experiences of a new-man-in-the-field as he tries to systematize the data he is collecting. They are inserted, apparently, as an emphatic reminder that each scholar must learn for himself the means for bringing order out of confusion and must face up to the inevitable stage when he becomes swamped with detail. Alternative fact-finding devices—the questionnaire, tapping the ongoing information-gathering services, and the interview outline—are each taken up at least once. It is apparent, however, that anyone coming back to the scene of an original study, armed only with such eye-witness reports as are reported here to serve as the basis for his expectations, is in for some serious surprises. These studies, and the methodologies for which they serve as illustrations, are much too partial and tentative to be used as foundations for simulation.

In contrast, another collection, labeled *Comparative Studies in Administration* (Thompson, Hammond, Hawkes, Junker, & Tuden, 1959), brings together a much more carefully selected set of articles highlighted by several significant theoretical propositions. It proceeds from an analysis of the sources of differentiation between identical structures (two Air Force wings at the same base under the same commander using the same equipment) where the skills of individuals in key slots are the principal causes, to the remarkable differences of approach of two English-speaking cultures in the handling of dry cargo ships, and then on to the influence of size upon the capacity of welfare units to administer social services. Subsequently, the organizational biases induced by the milieu are explored for (1) grant-supported research in an academic institute for social studies, (2) the manufacture of steel on two continents, and (3) central Algonkian patterns of life as maintained to the present. The process of exerting authority and making decisions in two different departments of a hospital, and in two similar-sized firms in Norway, are considered in the light of variables currently emphasized in management science.

This sampling of the administrative world avoids to a very large extent the description of ponderous bureaucracy and the clichés to be found in the manuals for regulating such structures, yet the deductions (when generalized sufficiently) are as relevant to large organizations as to those of medium size. Also, they are dedicated to the "steady state" that is only achieved after an organization has become established. The excitements and pathologies of initiation and rapid growth are eschewed.

It is worthwhile introducing at this point another current study, *Automation and the Worker* (Mann & Hoffman, 1960), which purports to analyze the influence of modern technology upon the human organization that operates it. It applies the techniques of social psychologists as of 1954–56 but these are readily seen to be far from adequate. For example, some relevant qualitative points are made about the elimination of dirty and tiresome jobs in the course of introducing "automation,"[1] which means that the pace of decision making by workers and supervisors must have fewer peaks and

[1] The most modern plant in the study was only relatively automatic. Power plant design has moved another step forward since then and "closed the loop" so that a computer makes most of the operating decisions. This means that the conclusions can not be expected to apply to the 1960's.

valleys. Nothing is said about the information exchange rate between men and equipment. Nor is there any mention of the accident rate,[2] although one might deduce (if he knew nothing about the development of fail-safe devices in power plant design) that errors would be more expensive in the highly capitalized modern plant than in its predecessors. The material provided on population density in the plant and communications between work positions is completely inadequate. The description of social structure and its evolution, however, is fairly thorough and illuminates some features of shift work which are not widely understood. There is only an odd bit relevant to planning horizon, while product competition and substitution are rightfully excluded as irrelevant. To summarize, a much more comprehensive analysis than this would have to be made if we are to understand the influence of even one technology upon human organization. More than that, a future study should include an industrial engineer who could, in conjunction with psychologists and sociologists, consider the range of choice for organization structures that do not add significantly to unit costs of production.

A new monograph, *Organizational Systems and Engineering Groups* (Barnes, 1960), has provided us with a much more useful study of conditions in the real world. This author has considered two similar parts of different organizations as if they were systems, one of them chosen because it is relatively closed as judged in terms of contacts with other sections of the parent organization and the other relatively open. The latter is seen as not necessarily more productive (the measures of productivity of engineering groups would be rather particularistic and arbitrary) but much more flexible and innovative and therefore more useful for any organization committed to growth.

Barnes' contribution is a fascinating combination of observation and theory which tells a great deal about the range of variation in the organization of professional technical talent. Contemporary general theory as promulgated by Merton, Homans, Roethlisberger, and the group dynamics circle is used in explaining behavior and suggesting hypotheses. Thus a relatively complete picture is obtained of the interaction between external relations of the group and its internal structure and work procedures. If it were not that the informational requirements and outputs of an engineering group engaged in product improvement and trouble-shooting were so complex, we would have sufficient grounds for simulating a department of a large organization. We obtain from Barnes' effort two models for achieving co-operation, but for each of them altogether different conditions are assumed; one of them is relatively insulated from outside events and crises, while the other is very much in tune, considering the events as opportunities and challenges. The facts in this instance have undergone several stages of interpretation and evaluation. Another difficulty that discourages simulation, based upon Barnes' studies, is that the subset of which his cases are examples can be expanded only to several hundred thousand persons (if all technical service and engineering design groups were to be included), or much less than one per cent of the total labor force. The returns to be anticipated from considerable effort undertaken in this direction are therefore not particularly great, because the supply of engineers filling these roles does not promise to become a stumbling block.

In the book *Understanding Organizational Behavior*, Chris Argyris (1960) presents us with a carefully constructed how-to-do-it volume which depends heavily on face-to-face interviewing techniques that can be employed for understanding production-oriented organizations. The underlying theory, and the specific model of social organization, with which he works is one that follows naturally from the source of his information. When Argyris diagrams his social system, it is not built up from people, or even roles within which they operate, but from a set of social process phenomena and related abstractions. The components of the system can be labeled as values,

[2] Changes in accident frequency can be a very significant variable in a well-organized plant. (See Haberstroh, 1960.)

attitudes, concepts, procedures, conditions, policies, and strategies, and they are linked spatially in a paradigm. The interview questions, and the techniques of interviewing, as well as the scale of self-actualization employed for rating the protocols, are admirably well-developed and handled in a critical, self-aware fashion. Yet each paragraph in the book is understandable, even to the serious business manager! This effort is presented as an interim report, with the potential of the ideas and methods incorporated in it still incompletely realized, so it gives the impression of being a humble document. Argyris already knows that he can find out things about a factory organization that could not be discovered by the management or the normal run of consultants that could be hired to make a study. He is able to penetrate to some degree a substratum of invisible attitudes and relationships in workers, foremen, and managers that are precipitated by organizational processes. This kind of information is invaluable when it is necessary to understand the impact of innovations upon factors affecting worker and plant productivity. However, this work provides only a partial description of the whole organization and its behavior and bears little resemblance to the standard operations research kind of study.

Argyris and Barnes both emphasize the degree of their dependence upon earlier work on the study of human groups as manipulated by leaders, administrators, managers, military officers, and experiments. Society has been its own laboratory for thousands of years, but only recently have the sociologists and social psychologists begun to sort out the variables. The literature is now replete with an astonishingly complex and confusing series of evaluations and experiments. In *Leadership, Psychology, and Organizational Behavior* (Bass, 1960), we are provided with a most useful analysis and review of this literature. This volume covers the contents of 1155 citations within the space of 648 pages, but unfortunately it is already obsolescent because it goes no further than 1955–56. One redeeming feature is that it draws upon history and the classics for many examples of noteworthy relationships, but not at all in the humanistic tone affected by some psychiatric writers. This is a no-nonsense mining of history for suitable illustrations of situations that are otherwise treated by contemporary observation and experiment. The "organization" in all these studies, however, remains simply at the level of relationships between the leaders and the led, and those discovered among the led in the various situations that they may find themselves. The upper reaches of hierarchy are never probed.

ESTABLISHING THE SYSTEM

Anyone who attempts to recapitulate a sequence of social events quickly loses his awe of the hugeness of the computing capacity now available to him. He is impressed instead with the limitations placed upon his effort by the actual capacity of the equipment. Only a few purposes can be served; the others are too prodigal of machine time. Costs can build up very rapidly and still no significant results are obtained. Therefore every simulation must be concerned with "getting to the heart of the matter" with minimal drain upon the equipment and its operators. We are on the lookout for new techniques of economizing which will appear as convenient methods of determining boundaries of systems, differentiating between levels of organization, detecting redundancy, and identifying relationships.

The best example of the kind of field work likely to generate the prerequisite data was actually loosely associated with some of the pioneers in the simulation of social institutions. In an article in *Human Organization Research*, Allan R. Holmberg (1960) reviews the methods employed in an action research program on Culture and Applied Science originating at Cornell University. The technique of these researchers was to rent a hacienda in the high country of Peru and then do their utmost to introduce constructive change, making sure that this change was institutionalized firmly enough so that the innovations were likely to survive and develop further after these *Norte Americano patrones* had left. The program for change was conducted like a

sequential study, with each new unanticipated finding leading to new hypotheses, new goals for the community, and new plans for probes, protests, reactions, interventions, and appraisals.

The same approach was claimed by John L. Kennedy (1960) in a succeeding paper entitled "A 'Transition-Model' Laboratory for Research on Cultural Change." The term "transition model" is borrowed from the intermediate stage in the development of aircraft when scale models are subjected to wind tunnel tests. This paper describes a "cultural wind tunnel" for a unit of the air defense system. Asserting from experience in studies carried out by his group that "people behaved like people only when the environment was complex, rich, and challenging," they took five officers and thirty-five men for each of four model organizations. College students, they discovered, did not behave like the military and therefore their performance could not be extrapolated to the large institutional setting. This paper stands out as a classic for those who are interested in the synthesis of human organization.

Perhaps their most significant finding is the need for running through a varied, nonrepetitive task many times. In another place, Milton G. Weiner (1960) has summarized the experiences of persons involved in these simulation exercises. During a *basal* period, he notes, the participants attempt to transfer their previous experiences in perception to the new task. This is followed by a *consolidation period* during which previously learned strategies are applied to see whether they would lead to improvement in performance. After this selection process has proceeded almost to completion, and inefficient strategies have been excluded, a sequence of *organizational invention* may take place when altogether new combinations of policies and reallocations of effort can appear. This assumes, of course, that the organizational task is a demanding one and that the "transition model" in the research and development process (in our terms, the simulation of a social institution) is equally demanding.

At the School of Business of the University of Chicago an altogether different procedure was employed for seeking out systematic definitions for synthesized organization. Recognizing the truism that is repeatedly cited by the sophisticated computer salesmen (low pressure type), namely, that the prime function of a computer in an ongoing organization is to stimulate systems analysis and then provide the capability for taking advantage of both the new data and the fruitful insights, they gave Simon and Newell a challenging topic and sought out the industries with computer experience that were willing to respond. In *Management Organization and the Computer*, Schultz and Whisler (1960) have recorded the results of a seminar on the theory and practice of management of organization as modified by the presence of one or more computers. Simon and Newell here extend my argument; they feel that within two decades computers should have enough capacity to take over the duties of most middle level and higher executives. This could only be done, of course, if heuristic programming and similar advanced techniques have developed further. They make the point, however, that the cost is not inconsiderable and therefore the full-scale simulation of an executive would be undertaken only in those instances where executives could not be obtained more cheaply. The case studies that followed led to the suggestion that there is likely to be some stubborn resistance in the rest of the organization to an electronic executive; the total costs are substantially greater than those of operating and amortizing the big machines.

The evaluations of the respective adaptations to the introduction of the computer are of mixed quality. The International Shoe Company had been moving into a morass, but apparently grasped this chance to recoup. It discovered that its central problem was the management of inventories of perishable products.[3] United Air Lines found that the crucial questions for them

[3] This industry has been particularly fortunate because a modern analysis by Ruth P. Mack was followed by two independent simulations by K. Cohen and J. V. Yance (quoted from M. Shubik, 1960).

had to do with variations in the routing and maintaining of aircraft. They proceeded to construct an ambitious simulation of an airline maintenance facility which clarified certain strategies and was then used for educational purposes in maintenance operations. A working model of the flow of information through a reservations system was also constructed. This is a firm from which much can be learned. At Esso, on the other hand, astonishing quantities of computers and electronic data processing were gulped down and digested without noticeable transformation. A period of rapid change may yet come, however, because at the world headquarters in Rockefeller Center the petroleum statistics, financial reports, and personnel data have been reduced to electronic procedures. The revolution might well follow the next major international shock, when it is realized that a thoroughly compartmentized organization responds too sluggishly. The Atwood Vacuum Machine Company (a nongiant with 1500 employees, located in Rockford, Ill.) more or less blundered into electronic data processing without the aforementioned systems analysis. This firm profited the most from relatively standard methods of inventory and accounting control which permitted it to operate more as it did when it was quite small. This added data processing capability enabled the founders, who are still in charge, to run a relatively large organization in the style to which they had become accustomed.

THE CONSTRUCTION OF SIMULATIONS

One of the most ambitious simulations brought off to date is that of Guy H. Orcutt and collaborators (Orcutt, 1960; Orcutt, Greenberger, Korbel, & Rivlin, 1961; Orcutt, 1957). They started from an econometric model by Lawrence Klein, and incorporated a representative sample of simulated households which formed, grew, diminished, split, and disappeared according to statistics that fit the past, present, and expectations for the immediate future. Thus two "blocks," one of them demographic and the other economic, have been integrated into a program in such a fashion that estimates of aggregate population and income are generated. As time goes on, new blocks can be added and the number of parameters may thereby be reduced at the same time that more detailed projections are achieved. An important advantage is the ease with which the subprograms can be improved and adjusted with new findings from survey research and new economic series. This is the first simulation outside of the military sector which has been designed to meet the needs of policy makers in government.

In a recent article, Clarkson and Simon (1960) provide several unusual examples of fully programmed simulations of complex tasks, such as (1) balancing an assembly line, and (2) selecting an investment portfolio, which are taken all the way from field studies to the complete computer simulation. These are really psychological simulations because they portray the train of logical processes required for a respectable degree of success by a specialist in an organizational context.

Conferences on simulation techniques are held quite frequently these days, sometimes as a convenience for the proponents to spread their gospel (yes, there is already a circle of enthusiasts with missionary zeal), sometimes as a feature of the reorganization of a revamped professional school, institute, or department, but this time an attempt to secure contact between studies initiated under the cloak of military secrecy and those in private corporations where the urge for concealment is less well implemented yields a few suggestive points. A conference on Management Information and Control Systems was called by the System Development Corporation of Santa Monica, where most of the simulation preceding the synthesis of the SAGE operations and some successor systems was carried out. The contributions in *Management Control Systems* (Malcolm & Rowe, 1960) are qualitative but at times insightful, and they bear upon the experience of the authors and organizations with simulations procedures over a span of about seven years. Thus, much of what they describe now must represent the survival of the fittest models among all of those that could have been invented.

Rathe's paper in this volume is par-

ticularly useful for educational purposes because, within an appropriate context, it contains several checklists relating to management controls in business. It reaches one high point when tabulating "57 varieties" of control and another when evaluating the functions of the ubiquitous corporation controller.

Integrated systems planning at General Electric reveals the same approach to the reorganization of social entities that was proposed by Holmberg, Kennedy, and Weiner earlier, but an extra feature is consciously built into the system. Ford Dickie introduces the principle of "synergism," which economists will immediately recognize as Marshall's "external economies" after the system description has been translated into economic jargon. The search for out-of-the-ordinary interactions between elements, either positive or negative, has a good chance of leading to surprising improvements, some of which are likely to have revolutionary implications. (Perhaps some of Simon and Newell's heuristics fall into this class, although so far they apply only to the programming and not to the underlying system.)

Because the book went through the publishing process with surprising celerity, the discussion of the organization principles in the "second generation" business games still sounds new. These games now include price-setting decisions, production scheduling, hiring and firing, cash flow, co-ordination of departments, reorganization, etc., in a simulation free of human intervention. The reduction to underlying routine should make executive behavior when playing these games quite a bit more realistic. Although the following quotation is still marred by jargon, any player of management games will note the similarity of the principles expressed to those espoused by Orcutt for the national economy:

> "These first efforts at the simulation of a complete business firm have pointed out the difficulties and obstacles to simulation procedures. However, they have proved reasonably successful and the Task Manufacturing Corporation is undergoing a major overhaul in those models already completed. Initially, the amount of detail included was too great. The production process was run daily with the machine model run hourly and the employee model daily. Even for the limited number of models in the simulation, computer time rapidly approaches an upper limit. The first change will be designed to make models run for either one week or one thirteen-week quarter. It appears that this can be accomplished with only parameter changes.
>
> "As additional models are added the Task Manufacturing Corporation will simulate a complete business firm. All the subsystems listed earlier will be added.
>
> "As efforts in simulating the complete firm continue, we envision many models of component subsystems. Eventually there would be a library of such models available to serve as the building blocks from which a particular system simulating a particular firm may be synthesized. Some of these models may reflect in great detail a theory about the behavior of their real-life counterparts and be useful for testing hypotheses about its behavior. Others may be merely the sampling from distribution functions which appear to reflect real-life data and behavior without any attempt at modeling a theory behind the phenomenon being simulated. Further effort and experience is needed to test an original premise that the simulation of the operations of a complete business firm is a fruitful educational and research undertaking." (Sprowls & Asimow, 1960, p. 332).

The most recent and most thorough of the reviews on management games (Cohen & Rhenman, 1961) reports only one published study that makes some contribution to theory (Hoggatt, 1959), in the sense that the term is traditionally employed, but virtually all are intended to facilitate teaching and training.

Cohen and Rhenman's review is a landmark. It carefully resurrects the history and experience with antecedents to management games from obscure works, unpublished papers, and technological reports, and then describes in essential detail the evolution of the major games now in existence. Their discussion concentrates upon the Carnegie Institute of Technology game used for graduate instruction in industrial administration, perhaps the most complex of all because several hundred decisions are required on the part of each team each week. It led the authors to outline a few first steps toward establishing what they believe will become a systematic body of knowledge. The argument is sober and persuasive. The various possibilities that can now be explored within the context of "management"

are considered and then optimism is expressed regarding the "dynamic case study" which may move in such various directions as negotiating a labor contract, arranging a loan, or managing public relations in addition to the more standard inventory control, advertising outlay, research and development expenditure, etc. Cohen and Rhenman insist that a management game differs from computer simulation and operational gaming (which usually employs Monte Carlo techniques for discovering the optimal solution for a game) because, in the first, no limits are placed upon the strategy; in operational gaming, they see a future primarily in the development of organization theory.

THE EXPOSITION OF THEORY

Fortunately, in a book called *Some Theories of Organization* (Rubenstein & Haberstroh, 1960), we have now a cross-disciplinary hybrid between a reader and a text that is suitable for introducing the honors student and first year graduate student into alternative and overlapping views on the theory of organization. The reprinted materials were authored by many of the famous names, but also include a score or so insightful pieces that would otherwise be lost. The editors chose to subdivide the materials into (1) the nature of organization theory, (2) organizational structure and process, (3) leadership and morale, (4) communication, (5) control and evaluation, (6) decision making, and (7) research techniques. As is usual, each of the editors includes a piece of his own in the collection, but it is not so common that they stand out so well in comparison with the gems of the past output. Rubenstein's paper on measuring communication in industrial research clarifies organizational interaction in a lead sector (e.g., one which sets a pace and is likely to influence other sectors so that they behave in a similar manner). Haberstroh's article on error frequency (accident rate) as a function of innovation rate was referred to earlier. It offers a prime justification for applying the engineer's control theory to the understanding of human organizations, and contributes a very nice approach to the estimation of some social costs of innovation and adaptation which were hitherto invisible. The editors have done one other thing to put this book into the textbook class—they include a series of probing questions in an appendix which assist in the comprehension of materials that fall outside of one's own area of competence.

Organization theory is a slippery, tantalizing subject. Promoters and entrepreneurs who operate intuitively seem to create organization out of near chaos more often than they fail. Indeed, at the psychological level, the rankest amateurs—ordinary infants—create an astonishingly good organization of concepts as they learn about their environment. Yet, as scientists, we are unable to draw up a blueprint for a viable organization which identifies the respective components, both human and mechanical, and have any confidence that the unit will work as designed. The military puts on a show of synthesizing task forces in just such a fashion, but everyone knows that the real work of adaptation gets done through informal contacts and private arrangements which cannot be mapped out in advance. The formal structure imposed upon a military group is often a real hindrance except in an emergency. Perhaps the best that we can do at present is to convey the principles by means of practice, using exercises in collective action which appear to convey concepts of organization even before they have been completely isolated and named.

The games, exercises, and tasks involved in the simulation of organizations are contrived behaviors that have been abstracted from real life. Such behavior is enacted out of context, but the sequences of action and decisions should be reminiscent of reality; they create a context of their own which is not incompatible with what is being modeled. They become important when the exercises *precede* personal or group experience, because the decisions made later in real life could easily be influenced by those in the game. The outcomes, if perceived as interesting or desirable, may serve as predictions of subsequent choices. A good experimental game or task, it is agreed, can be as spare and stark as chess or a Greek

drama, or it can be as rich in details and background as a novel or a day in the mayor's chair.

The methodology for creating satisfying simulation involving human participation has evolved to the point where it can be outlined:

1. Identification of a salient problem. While time and skill remain scarce it is preferable that such problems be chosen from among that portion which are likely to be of increasing importance in the foreseeable future.

2. Completion of thoroughgoing case studies which isolate the system, the major variables, and the key decision points.

3. Demonstration that the dynamic system cannot be adequately described by simple logical or mathematical formulations; in other words, it must be shown that the relationships are not simple, but complex, and the outcomes are unpredictable.

4. A team must be formed, combining practitioner, researcher, and programmer, which organizes elements of the exercise for one of the following purposes: training of personnel, measurement of human capability, decision making (i.e., exploring the range of alternatives), research in group processes, the display of interactions within a complex system, and therapy or recreation. Sometimes secondary purposes enter into the design.

5. The program is "debugged" in the course of trials, and the "rules of the game" are adjusted so as to remove defects and make progress toward the self assigned ends.

6. Then standard players are introduced into the roles so that the scores obtained are comparable and can be correlated with other individual or group characteristics, thus connecting the findings with the main body of psychology, industrial engineering, administration, etc.

7. The successful result is embodied in a curriculum, training program, sequence of investigations, or therapeutic program which exploits its potentials for creating organization.

Serious gaps can be detected in the repertory of simulations involving human participation that have been developed to date. The majority now in existence (approximately 100) emphasize improved calculations of private profit (corporation executive games, management games, marketing games, inventory control exercises, etc.). An important fraction is based upon military conflict situations employing attrition and survival in combat as central considerations and may therefore be regarded as descendants of chess and other military games. A few examples are based upon popularity polls within simulated voting systems. Far too few formulations are available which dramatize and encapsulate situations where the players are expected to discover the *public* interest and are scored according to their degree of success.

Yet the exceptions to a generalization about lack of attention to public needs are as illuminating as any of the above, if not more so. Orcutt's computer simulation of the American economy does not strictly belong to this class, but could probably be operated with human intervention. Perhaps the most intriguing of these is the Northwestern University Inter-Nation game (Guetzkow, 1959). In it, five to seven nations, played by three or more "office-holders" apiece make internal economic and military decisions and engage in trade and diplomacy over a dozen time slices that represent budget periods. After three years of development, it is playable by high school students (Guetzkow, Brody, Driver, & Beach, 1960) and shows that small differences in prior opinion in the players can be magnified by interaction between these "nations." For example, the game that contains more players holding a belief in *real-politik* explodes into atomic warfare not long after small countries obtain nuclear weapons, while the game with a higher proportion of players believing in the limitation of sovereignty and conciliatory behavior on the international scene achieves "international stability" despite the threat. The results are no more predictive of the behavior in the real world, given these eventualities, than is a mathematical model, but the number of insights to be derived from a study of the process is much greater. The

suggestiveness is obtained at the price of some ambiguity in the interpretation of the experimental observations.

THE PATH OF VULGARIZATION

The inventory of available simulations needs elaboration and extension in this and other directions. The role-playing was originally performed by industrial managers and military officers, but now the players are more likely to be apprentice managers, non-commissioned officers, and new graduate students. Some management games are regularly played by undergraduates. There seems to be no reason why the vulgarization of some of these models cannot continue so that some of them will soon appear in a high school environment supervised by teachers rather than university investigators. Then "the game" can be used to give spice and flavor to rather ineffectual attempts at explaining how political, social, and cultural systems actually work. The result should be the eradication of many misunderstandings now held about our social institutions and the constraints that they place upon public behavior. High school students have few settled preconceptions as compared to managers and officers, who tend to reproduce what they have already learned on a new stage, and so may acquire a great number of insights from the exercises.

The social structure of high schools in the United States has been shown to be biased heavily in the direction of athletic games. The leaders within the student body are attracted to sports because this is a means of demonstrating personal commitment to collective goals (Coleman, 1960a). However, these skills are virtually irrelevant in adult activities, and so constitute a diversion that wastes much effort that could have been devoted to the development of talents for effective leadership. Good quality games employing more substantive materials, such as literature, history, law, logic, etc., should fill the gap (Coleman, 1960b). These devices have been considered before by educators, and employed in classroom work, but extensive judging was required, and once away from the faculty-supervised sessions in a laboratory school the judgments often became arbitrary and small-minded, so that really advanced performances on the part of the students were inhibited. The simulations arising from the the extended research-and-development approach described earlier promises to reduce the amount of judgment demanded of referees, judges, and amateur impresarios. They also tend to require computing time in order to produce outcomes of actions and interactions with little delay. Thus far, there is hardly a high school that has access to a computer, but there are moves afoot to link a few pioneers among them to a computer installation by private wire in order to take advantage of new developments in teaching machines and closed circuit television.

Nevertheless, computing requirements, if they are extensive, will always stand in the way of large scale application in teaching situations. Therefore, for key social institutions, the simulations should start with simple group exercises where the play can be followed on a map or a chart and a reduced amount of computation is built into the scoring system and the display system. After the players have learned the elementary principles and are ready to proceed to a higher level of abstraction, some parameters can be allowed to become variables, so that more complex strategies can be evolved. Then at the next stage, perhaps in junior college, they would be ready for the "third generation" exercises using extensive computer-operated subprograms.

If simulation exercises are to be successful in the secondary school environment, the outcomes must be easily apprehended, the play must be open enough to be followed by an audience, and there should be many discriminable levels of skill. Basketball has been synthesized as a game over the last half-century, and the scores in the professional games show that there is still evidence of steady improvement in skill. It provides us with the nearest analog for the selection of suitable games. Eventually we may expect that part of the basketball and football contingent in high school could be deflected into equally intriguing activities which here are referred to as simulations

but would, of course, have different names. Some high schools may specialize, because they have a good "coach," while the student bodies of the larger ones would be likely to subdivide their attention between several intriguing games.

One could expect also new bases for college recruiting. At the moment, colleges recognize that "something else" is involved in success at college than athletic and academic prowess. Thus far, there are few indices of the capacity for leadership other than participation in social activities and general popularity that appear on a student's secondary school record. The acquisition of skills relevant to the operation of modern social, political, and cultural institutions should be most helpful in discriminating special aptitudes that now are often overlooked and constitute an important share of our under-used human resources (Wolfle, 1960).

NEW DIRECTIONS

It behooves one who criticizes the present state of affairs to be explicit about the kinds of developments that would, in his opinion, constitute marked improvements. Thus the games or exercises that might be used for introducing an elementary understanding of basic social institutions ought to be described in a preliminary fashion, and some of the criteria for assessing the quality of the simulation should be specified.

If we are to argue for subcomputer size simulations, and make them simple enough for use at the secondary school level, it seems likely that each must be designed to develop several educational objectives at once, depending upon what the student brings to it. Referring again to the basketball parallel, such a gymnasium game serves many players as a convenient source of exercise, but for an important number it teaches the principles of finely co-ordinated teamwork. In general, the kinds of principles that are taught will depend upon whether the student is a member of a class exposed once or twice to a simulation or an enthusiast choosing to spend his spare time with it. The following outlines suggest complementary directions that such simulations might develop:

1. The structure of language and the theory of the firm. The micro- and macrostructure of written language, as presented in various modern classics, offers an excellent basis for gaming. Following some experiments by Clark and Ackoff (1959), it seems quite feasible to set up groups which buy random batches of letters, manufacture them into words, and "sell" these words on a market, with the option of buying selected vowels or consonants from "middlemen" at higher prices. If we make the "market" the wording of Milton's sonnets, the Constitution of the United States, a selection from a reference work, etc., there is considerable assurance that such texts will be well-remembered even though their significance is not yet well understood. The possibilities are much superior to spelldowns, junior businessmen's projects, classroom memorization contests, and similar devices for building motivation. The text being used can be displayed on a scoreboard so that nonplayers may follow the progress of the competition, and they, too, acquire a familiarity with the text.

2. The synthesis of organized networks. The networks of communications, movements, responsibilities, and the like that are created by highly organized systems have different properties than other purely mechanical networks. Thus the connectivity of a network placed upon an abstract map or grid may be used to represent increasing degrees of organization. Teams may build networks, receiving resources proportional to its degree of "organization," but using them either to hold back the opponents or to grow themselves. The problem is to create a viable, relatively invulnerable network and see the spatial pattern evolve. Since all organizations seem to be reducible upon analysis to a set of positions and relations reacting to a flow of events, i.e., a dynamic network, the properties of many different kinds of organizations can be represented by alternate initial conditions and rules for the use of resources.

3. The reproduction of history. Significant eras of history can be replayed, whereby

the external events influencing a set of decision-making centers may be introduced randomly, or in different sequences, so that alternative possibilities for development come to be understood by the players. The wealth of dramatic presentations of the settlement of the American western frontier transmitted by television provides a source of hypotheses (not all of them inaccurate). A simulation of this sort can be used to restore a more balanced view of past events.

4. Public service operation. A school administration, a library, or even a river basin, may be reorganized to fit the needs of a public. In this instance, various kinds of tests must be undertaken in an uncertain environment in attempts to discover whether an innovation will pay off. The adoption of an innovation, if not disastrous, would generate increased inflow of resources, part of which could be invested in further innovations. This is a management game also, but the returns are measured in estimates of public welfare and not in profit to the organization, although the organization is likely to set its own survival as the goal having highest priority.

5. Community administration and planning. Problems of bargaining between special interest groups on the local scene and their implications for the public over the long run can be illuminated. Professor Nathan Grundstein (1960) at the University of Pittsburgh has been pioneering in this direction. A mayoralty and city council election may serve as a subsidiary exercise in this context.

6. Automotive fleet simulation. The equipment used for teaching automobile driving can be linked with a schedule of costs for moving about in a metropolitan area. The simulation here could take advantage of mathematical experience with the "traveling salesman problem" whereby various addresses appearing at random are contacted by any one unit of a fleet while minimizing cost. The atmosphere of such a game can be designed to attract nonintellectuals and introduce them gradually to very respectable levels of problem solving without intimidating them. Once these youngsters discover that they have aptitudes respected by society at large their aspirations suddenly expand, so that they may then work up to highly responsible positions in society.

7. Population dynamics and ecology. The survival of competitive species in a risky environment can be played out as a representation of wild animal populations. As a part of the nature studies being incorporated in many curricula, this simulation would generate more realistic experience than could be acquired by field trips. A similar, but more complex, game might be prepared which is based upon the properties of human populations which might eventually lead to a better informed national population policy.

The foregoing suggestions represent only a limited sample of the directions such efforts may take. With very little trouble a fertile imagination could, when followed by a series of small side experiments, generate an equal number of others with as wide a range of properties.

There is a general property of games which has been treated intuitively but recently has been asserted to be readily quantifiable. In a note studded with sly humor and some obvious puns, a British statistician has presented us with a simple criterion for assessing the "goodness" of a game (Good, 1960). The population of learners and trainers of a game will distribute themselves in a way such that grades of skill can be identified. A very convenient basis for the separation of these grades is an ability to defeat members of the next lower grade two out of three times. The number of such grades serves as a measure of the degree to which the game requires skill for success. The concept of *difficulty*, in the common mathematical sense, is quite different because it applies to puzzles or theorems which have at most a few successful pathways, while the goodness of a game depends upon many pathways which have predictably different efficiencies under the playing conditions.

The number of gammas in chess, by such a definition, would be very high, the game of GO would also be high, while checkers would be markedly less. All of these games would have greater goodness than most of the

commercial parlor games that claim to simulate situations in life. Athletic games based upon team play have evolved over time to levels of goodness approximating those of chess, with the top professionals being as carefully selected as the chess masters. We have evidence from tournaments and round-robin seasonal play that leagues of competitors with three gammas retain our interest, but when it expands to five or beyond the competitions between top teams and the cellar-dwellers are too one-sided, with the result that graded leagues then begin to evolve. Any game suitable for high schools should be able to differentiate teams of casual participants, regular but relatively mediocre performers, and highly practised star performers who could compete successfully as representatives of the school. Therefore, a suitable game must have a minimum goodness index of 10 gammas, again a figure that exceeds most parlor games.

Once the minimum goodness has been achieved in the design of a game, the strategy for further development diverges. The simulation can add descriptive detail so that it resembles more closely the social institution after which it is modeled. Few tasks in the real world differentiate excellence to a degree equivalent to 10 gammas. Luck, style, and various personal qualities still have considerable weight, and in the drive for realism it is proper that they enter the picture once the minimum conditions for stable competition in the game have been achieved. A few games with very high levels of goodness might also be introduced, particularly to assist in the mental development of the top one per cent of the student body. Little else in a high school curriculum is likely to be as challenging as a good quality game resembling closely some feature of the real world. Thus a battery of simulations is required whose characteristics are distributed according to two known dimensions— empirical detail and degree of abstraction achievable.

SOCIAL CHANGE FROM SIMULATION

The existence of a need for simulations of social institutions at the subcomputer level for purposes of unifying the secondary school curriculum does not mean that these exercises will promptly come into being, but let us suppose that a dozen or more working in much the same manner as envisaged above do appear. Their diffusion through the high schools might well parallel the introduction of teaching machines and the "programmed textbooks" that are likely to be the newest educational fads. The transition could take no less than a decade, and could easily require two. If simulations become popular, what consequences may be expected?

First, if present descriptions of recruitment of talent in American secondary schools are correct, it seems quite likely that human resources, particularly intelligence, would be more efficiently exploited. With added relevant scores upon which judgments could be made, an increasing fraction of each cohort would qualify for professional training of one sort or another. The degree of variation in family background and personality would be extended somewhat because the increased utilization would be drawn heavily from lower income groups, minority racial groups, immigrant families, and the like. The extra component would also be more female than male. It is within these categories that most of our wasted talent may be found. The variety in backgrounds promises to open up a wider range of organizational structures and permits more choice in geographical location where new organizations could succeed. This trend is expected to be so gradual, however, it would appear to be small as compared to other sources of social change, especially those emanating from innovations in technology.

Second, experience on the part of high school and college students with different kinds of institutional prototypes should simplify the task of coming to agreement upon what forms of organization would be appropriate for a given new task. These students, when they have matured, should be able to form new organizational structures quickly, learn rapidly as a group, and dissolve and reform into new organizations with markedly greater fluidity. A single individual under these circumstances would

find it possible (probably even expected of him) to be a member of many more organizations in a lifetime than at present because the necessary mistakes and fumbling were made earlier in life. Also, various members of society will have shared these common experiences so that the precision of communication is enhanced.

Third, it is possible to take some of the outputs of one simulation and make them the inputs to another. The formation of backward and forward linkages between two systems can then be observed under laboratory circumstances. To be explicit, the repercussions of decisions in a "public utility" game may be used to modify the environment incorporated in a "population dynamics and ecology" game. Then a policy for maximizing the amount of water available for irrigation and hydroelectric power has one consequence for wildlife populations, and a policy favoring flood control has another. At present, the linkages interpreted through human judgment are poorly understood so that the existence of a battery of well-tested games may lead to a more thorough review of the possible outcomes from a given action than was ever possible before. We seem to regard the real world as an assemblage of loosely-linked systems, so it is to be expected that this view will result eventually in the setting up of complex experiments that would be too expensive to carry out full scale.

Fourth, if ever a basic theory of adaptive systems is produced at the mathematical level,[4] then the battery of simulations would be designed to convey the basic concepts so that they could be readily apprehended as intuitive insights derived from experience. Then the simulations could be as carefully standardized for the schools as the achievement tests. The discovery of a general theory of adaptive systems would hasten the processes described in my first three points and would very likely carry social organization to levels that are not credible in terms of contemporary values. We cannot afford to refrain from this change for fear of unknown repercussions because our present values with respect to national sovereignty, resource use, and family size, to take three outstanding examples, we are quite certain cannot be retained for many more decades without destroying civilization. A much higher degree of flexibility in social institutions is needed than has ever been achieved in the past, and in simulations we have one of the most promising possibilities for acquiring relevant experience at low cost.

Thus, still rather faintly and without too much confidence, the potentials of a technique can be seen which would increase the adaptive quality of a social system—a condition which most investigators interpret as being more highly organized. The resultant society would be able to meet a wider range of exigencies, and react with somewhat less confusion and delay than the best that we know.

[4] The recent works of both John H. Holland of the Institute for Science and Technology, and Merrill M. Flood of the Mental Health Research Institute, The University of Michigan, are especially promising. The insightful contributions of Oliver Selfridge of Massachusetts Institute of Technology and Sidney and Beatrice Rome of the System Development Corporation are highly suggestive.

REFERENCES

Adams, R. N., & Preiss, J. J. (Eds.) *Human organization research*. Homewood, Ill.: Dorsey Press, 1960.

Argyris, C. *Understanding organizational behavior*. Homewood, Ill.: Dorsey Press, 1960.

Barnes, L. B. *Organizational systems and engineering groups*. Cambridge, Mass.: Harvard University, Division of Research, Graduate School of Business Administration, 1960

Bass, B. M. *Leadership, psychology, and organizational behavior*. New York: Harper, 1960.

Clark, D. F., & Ackoff, R. J. A report of some organizational experiments. *Oper. Res.*, 1959, 7, 279–293.

Clarkson, G. P. E., & Simon, H. A. Simulation of individual and group behavior. *Amer. Econ. Rev.*, 1960, 50, 920–932.

Cohen, K. J., & Rhenman, E. The role of management games in education and research. *Mgmt. Sci.*, 1961, 7, 131–166.

Coleman, J. S. Adolescent subculture and academic achievement. *Amer. J. Sociol.*, 1960, 65, 337–347. (a)

Coleman, J. S. *Analysis of social structures and simulation of social processes with electronic computers*. Working Paper No. 4, Department of Social Relations, Johns Hopkins University, Baltimore 18, Md., 1960, 21 pp. mimeo. (b)

Good, I. J. The goodness of gammas. *Nature*, 1960, 186, 964.

Grundstein, N. Private communication concerning the Pittsburgh community model game. Mimeo, 1960.

Guetzkow, H. A use of simulation in the study of inter-nation relations. *Behav. Sci.*, 1959, 4, 183–191.

Guetzkow, H., with Brody, R. A., Driver, M. J., & Beach, P. F. An experiment on the Nth country problem through inter-nation simulation: Two case-study examples. Northwestern and Princeton Universities, Nov. 28, 1960. 55 pp. mimeo.

Haberstroh, C. J. Control as an organizational process. *Mgmt. Sci.*, 1960, 6, 165–171.

Hoggatt, A. C. An experimental business game. *Behav. Sci.*, 1959, 4, 192–203.

Holmberg, A. R. The research-and-development approach to change: Participant intervention in the field. In R. N. Adams & J. J. Preiss (Eds.), *Human organization research*. Homewood, Ill.: Dorsey Press, 1960. Pp. 76–89.

Kennedy, J. L. A "transition-model" laboratory for research on cultural change. In R. N. Adams & J. J. Preiss (Eds.), *Human organization research*. Homewood, Ill.: Dorsey Press, 1960. Pp. 316–323.

Malcolm, D. G., & Rowe, A. J. (Eds.) *Management control systems*. New York: John Wiley, 1960.

Mann, F. C., & Hoffman, L. R. *Automation and the worker—A study of social change in power plants*. New York: Holt, 1960.

Meier, R. L. Explorations in the realm of organization theory. I: Models of man. *Behav. Sci.*, 1958, 3, 68–78. (a)

Meier, R. L. Explorations in the realm of organization theory. II: The organization of space. *Behav. Sci.*, 1958, 3, 78–79. (b)

Meier, R. L. Explorations in the realm of organization theory. III: Decision making, planning, and the steady state. *Behav. Sci.*, 1959, 4, 235–244.

Orcutt, G. H. A new type of socio-economic system. *Rev. of Econ. and Statist.*, 1957, 39, 116–123.

Orcutt, G. H. Simulation of economic systems. *Amer. Econ. Rev.*, 1960, 50, 893–907.

Orcutt, G. H., Greenberger, M., Korbel, J., & Rivlin, A. *Microanalysis of socio-economic systems: A simulation study*. New York: Harper, 1961.

Rubenstein, A. H., & Haberstroh, C. J. (Eds.) *Some theories of organization*. Homewood, Ill.: Irwin-Dorsey, 1960.

Schultz, G. P., & Whisler, T. L. (Eds.) *Management organization and the computer*. Glencoe, Ill.: Free Press, 1960.

Shubik, M. Simulation of the industry and the firm. *Amer. Econ. Rev.*, 1960, 50, 908–919.

Sprowls, R. C., & Asimow, M. A computer simulated business firm. In D. G. Malcolm & A. J. Rowe (Eds.), *Management control systems*. New York: John Wiley, 1960. Pp. 321–332.

Thompson, J. D., Hammond, P. B., Hawkes, R. W. Junker, B. H., & Tuden, A. (Eds.) *Comparative studies in administration*. Pittsburgh: University of Pittsburgh Press, 1959.

Weiner, M. G. Observations on the growth of information-processing centers. In A. H. Rubenstein & C. J. Haberstroh (Eds.), *Some theories of organization*. Homewood, Ill.: Irwin-Dorsey, 1960. Pp. 147–156.

Wolfle, D. Diversity of talent. *Amer. Psychol.*, 1960, 15, 535–545.

(Manuscript received February 16, 1961.)

"It is very important for us that our goals shall in the main be unknown until we have achieved them; that they shall be larger than our means, and perhaps larger than our dreams; and, above all, that when they are finally within our grasp, we have the opportunity to pass on to the pursuit of other and more distant goals, still all but hidden over the horizon, still the foundation of dreams as well as of fulfillment."

CARYL P. HASKINS. *The Report of the President of Carnegie Institution of Washington*. 1959–1960

BOOK REVIEWS

Yves Galifret (Ed.). *Mechanisms of colour discrimination.* New York: Pergamon Press Inc., 1960. 296 pages.

An international symposium on mechanisms of color discrimination in man and animals, sponsored by the International Council of Scientific Unions, was held in Paris at the Collège de France July 25–29, 1958. The chairman, the renowned Professor Henri Pieron, at the time in his seventy-seventh year, must have provided a truly remarkable stimulus for the success of this conference. For a success it clearly was, as is evidenced by this report of the proceedings. This rather small volume brings together a summary of a good deal of the most recent research on problems of color vision, from experiments on the small compound invertebrate eye of bees and other insects to those on the much more complicated vertebrate eye, including—but not stressing—that of man, himself.

Anyone about to embark on an investigation of the color vision characteristics of any eye (vertebrate or invertebrate) can read with great profit the chairman's introductory remarks, which specify lucidly and in some detail the rather subtle problems and pitfalls in experiments of this type.

Newcomers to interdisciplinary research may read with some envy the discussions of this symposium made up of physiologists, zoologists, biochemists, psychologists, physicists, and ophthalmologists, all of whom seem to know just about as much about the other man's problems as they do about their own. Moreover, in vision, which has involved a mixture of disciplines from almost the very beginning, it is possible for the psychologist, for example, to understand the physicist's experiments and even to criticize them. Thus, frequently, the borderlines become blurred and one finds physiologists doing biochemistry, psychologists doing neurophysiology, and physicists doing psychology. The results of such a cross-breeding can be fruitful indeed and the evidence in this volume may serve as a stimulus to the interdisciplinarian who still struggles to understand what it is his "teammates" are doing.

But where has all this brought us in the problems of color vision as of July, 1958? Despite the enormous amount of study of the biochemistry of visual pigments over the last thirty years, the fact remains that ". . . no one can say with certainty that he has ever had a cone pigment in a test tube" (Dartnall). But we now know (thanks to the elegant experiments of Rushton) at least that such pigments do indeed exist. The evidence for the presence of two such substances in the living human fovea seems quite clear. The evidence is in this little volume for those who would relive the discovery, and it is one of the more exciting chapters in the physiology of color vision written in the last few years.

Also exciting are the discoveries coming to light as a result of the use of modern electrophysiological techniques on animal eyes. Thus microelectrodes impaled into single retinal (probably bipolar) cells of the teleost eye have indicated that equal energies of light can cause hyperpolarization or depolarization of such cells depending only upon the wave length of the stimulus! In single cells of the middle layers of the lateral geniculate nucleus of monkey, certain wave lengths of light, when stimulating the retina, cause a burst of activity at on—inhibition at off, while other (and complimentary) wave lengths cause the reverse (De Valois). In both of these cases the relation of these antagonistic wave lengths to one another is much what might be expected by Hering's opponent theory. The latter has been given additional impetus by the psychophysical experiments and theoretical formulations of Hurvich and Jameson, which give to an opponent theory an elegance one was accustomed to find only among trichromatic component theories.

The moment seems ripe for a certain synthesis of opponent (i.e., Hering) and component (i.e., Young-Helmholtz) color theories. Both Hurvich and Stiles, partisans for these two respective views, appear to be more or less committed to some form of unification of these two basically different ideas. Heated arguments will continue, of course, but perhaps more and more they will center around the stage at which emphasis should be shifted from a component to an opponent system, the relative importance of the two different systems, and the way they are linked together. This could very well be progress.

Electrophysiologically, in the lateral geniculate nucleus of the monkey the synthesis is geographical. In this case it appears that the two (one for each eye) most dorsal layers are arranged on a component basis and the two (one for each eye) intermediate layers on an opponent basis.

All this, and much more, is in this transcript of the Paris symposium. The interested reader can get a glimpse into these developments by reading this volume. But it will only be a tachistoscopic exposure through a keyhole, not a wide screen cinemascope view in three dimensions. The book is too small to find any of these matters elaborated in the detail necessary for their full appreciation. For that more extensive view, one must dig deeper, much deeper, into the original papers which are referred to in this volume. But narrow and brief as these glimpses are, they are none the less in full technicolor and some of the excitement is bound to rub off.

MATHEW ALPERN
The University of Michigan

Guru Datt. *Existentialism and Indian thought.* New York: Philosophical Library, 1960.

THIS book is divided into three sections, the first two sections covering 62 pages out of the total of 92. In the first two parts, the author has given a very systematic review of existentialism in Western thought; the third short section is devoted to drawing parallels with Indian thought. This section is poor in matter and presentation as compared with the first two sections. Existentialism in Western thought is discussed up to the present time but the source of Indian thought is taken mainly from the ancient Hindu scriptures with a collection of quotations from all sorts of ancient Hindu writings in an attempt to draw parallels with Western thought, sometimes these being far-fetched and forced ones. For example, on page 89, Mr. Datt quotes the character, Draupadi, from the epic *Mahabharata*, whom the Kauravs wanted to insult in the court by disrobing her because she was the wife of their rival brothers Pandavas. The author quotes this to draw a parallel with Kierkegaard's idea that existence will not stand publicness: "Existence is first and foremost, something which will not stand public exposure; that might indeed be the significance of the Mahabharata allegory of the failure of the attempt to disrobe the chaste Draupadi in the public gaze." In the context of the epic, Draupadi does not, by any stretch of the imagination, symbolize "existence."

Mr. Datt quotes Vedas, Upanishadas, Bhagvat Gita, and Tantras, in short every ancient Hindu writing he could lay his hands on. He refers to Buddhistic, Jain, and Sankhya schools of thought, Bhakti and Sakti movements, to emphasize the existential aspect of Indian thought in a very loose way. There are no footnotes given for his quotations and no bibliography at all. After reading this chapter, one is left with the impression that all Hindu writings emphasize some sort of existentialism. We wish the author had given more attention to the philosophies of medieval and modern thinkers, which contain a rich source of ideas on existentialism. For instance, the writings of the saints of the Bhakti movement, the Sikh scriptures, and the writings of Sri Avrbindo, Mahatma Gandhi, Rabinder Nath Tagore, Swami Rama Tiratha, and

Swami Vivekananda would provide the reader with some ideas which are quite close to existentialism. Addition of an index at the end of the book would have been very beneficial. The book does not contribute much to one's understanding of the existential aspect of Indian thought. But it does point the way toward the need for a more systematic study of the various systems of Indian thought with respect to existentialism, to bridge a gap between Eastern and Western philosophical movements.

R. K. JANMEJA SINGH and SANTOKH S. ANANT
Lafayette Clinic *The University of Michigan*
Detroit, Michigan

A good antidote to the monstrous proliferation of gobbledygook writing in government, business, and the professions is to wonder how the Bible would read if a revised version were entrusted to a committee of sociologists. Take, for instance, the opening lines of the Twenty-third Psalm:

> The Lord is my shepherd,
> I shall not want,
> He maketh me to lie down in green pastures,
> He leadeth me beside the still waters,
> He restoreth my soul.

The revised version would sound something like this:

> The Lord is my external-internal integrative mechanism,
> I shall not be deprived of gratifications for my viscerogenic hungers or my need-dispositions,
> He motivates me to orient myself towards a non-social object with affective significance,
> He positions me in a non-decisional situation,
> He maximizes my adjustment.

ALAN SIMPSON, "Liberal Education in a University"
Washington University Magazine

COMPUTERS IN BEHAVIORAL SCIENCE

Publication of this department is partially supported by a grant from the National Science Foundation to the chairman of its editorial committee, Steven G. Vandenberg

THE USE OF COMPUTERS FOR PSYCHOLOGICAL RESEARCH

A symposium sponsored by Division 19 of the American Psychological Association, and presented at that society's meeting on September 6, 1960, in Chicago, Illinois. Chairman: John E. Coulson, System Development Corporation, Santa Monica, California

Introduction: Recent Areas of Application,

John E. Coulson, System Development Corporation, Santa Monica, California

During the past several years psychologists and other social scientists have become increasingly aware of the potential applications of computers for research. At the risk of over-simplification, one might speak of three major areas of computer application in the social sciences: calculation, simulation, and stimulation. Many applications, of course, represent combinations of these areas.

Unquestionably the most common use of the computer in psychology thus far has been as a powerful *calculator* for the reduction and analysis of experimental data. Data analysis per se has been deliberately omitted from the present symposium, however. One reason is that computerized data analysis is a tremendously large and complex area, and could not be fairly represented in the brief period of this symposium. The other reason is that the problems involved in computerized data analysis are usually not unique to psychological research, but are generally applicable to the physical and life sciences as well.

A more recent and increasingly popular use of the computer is for *simulation*, particularly of human behavior. Theoretical models describing the behavior under investigation are programmed into the computer. Specific hypotheses, derived by the insertion of parametric values into the general models, can be verified by comparison with empirical data. Two of the papers in this symposium will deal with this application. Dr. Hare will describe a program for simulation of small-group interactions. Dr. McPhee will discuss his computer simulation of a much larger group—the American voting public.

This brings us to the third area of application: the use of a computer to *stimulate and interact* with human subjects. Computers may present complex stimulus patterns to an individual or group, and may also analyze the humans' response behavior. In Dr. White's research on perceptual processes, he uses a computer which generates visual patterns for presentation to human subjects. The computer-controlled teaching machine described by Dr. Silberman modifies its own stimulus outputs (instructional materials) as a consequence of certain aspects of the student's response behavior.

Among some social scientists, the feeling is occasionally expressed that computers serve no real function in psychological research except to add more glitter and prestige to the laboratory. For this reason I have suggested to each participant that brief mention be made, in the presentations, of the factors leading to the use of a computer in the research being described.

The Computer as a Pattern Generator for Perceptual Research,

Benjamin W. White, Massachusetts Institute of Technology, Lincoln Laboratory, Lexington, Massachusetts

For some years now, several of us at the Lincoln Laboratory* have devoted con-

* Operated with support from the U. S. Army, Navy, and Air Force.

siderable time to making movies on high speed, general purpose, digital computers. At first blush, this appears to be a frivolous way of employing the prodigious capabilities of such a device. Why use a computer to make animated cartoons?

The most obvious reason is that such equipment exists at the Laboratory, which since its inception has been engaged in computer design and the development of computer-driven visual displays. The cathode ray tube and camera required for this work are not standard computer equipment, and such units are costly and require considerable maintenance. The Laboratory not only has the equipment; it also has several people who are interested in visual pattern recognition, and in other areas of visual perception.

Let me give you an idea of the way in which patterns are generated on such machines. The computer can be instructed to display a point on the face of the cathode ray tube at prescribed X, Y coordinates. It can also be instructed to advance the film in the camera mounted over the scope face. The camera normally has its shutter open so that one may display a number of points in succession, the ensemble forming the pattern to be recorded. When this ensemble has been displayed, the film advances to the next frame, and the camera is ready to receive the next display. All of which sounds tedious until you learn that the computer can do this at great speed. One program I have run generates a display of 400 points, indexes the camera, computes new positions for all the points and displays them again at the rate of one every three seconds. The time required can be no shorter than the speed with which the camera can be indexed, which for our 16 millimeter camera, is $\frac{1}{10}$ sec.; and above this minimum time, the number of points to be displayed and the intricacy of the computations between displays govern the time required. At our installation there are a number of pre-programmed subroutines available to do such special purpose display tasks as drawing a line between any two points, or printing alphanumeric material in any location on the scope face either vertically or horizontally. Crude control over intensity is achieved by setting the scope brightness low and putting under program control the number of times each point in the display is to be "painted," and by putting brightness of the spot under control of the program.

These, then, are the display capabilities of the equipment. It is not just the computer's speed which makes it an interesting device with which to draw pictures; it is also the fact that it can perform so many operations on the coordinates which form the displays. A picture or pattern is a list of numbers stored in the computer, numbers which can be manipulated like any other numbers. All the rigid transformations are easily computed. Thus one may move the pattern around as one pleases, translate it on X and/or Y, expand it or shrink it, rotate it about the Z axis, or in three, or even four-dimensional space. Elastic deformations and nonrigid transformations are also easy to program. You will see immediately that here is an extremely flexible instrument for generating stimulus material for the study of the perception of movement.

Another useful feature of the computer is the fact that pseudo-random number subroutines are available. These make it possible to introduce controlled perturbations in a given pattern. Thus in a static pattern one can introduce a given amount of "noise" by randomly displaying or deleting points.

Some of the displays which we have made on the computer in this manner are shown in the accompanying figures. Figure 1 shows a dot matrix used by Green, Wolf, and White (1959) in their study of the detection of statistically defined patterns. In this display the matrix was divided into a number of strips. There were two probabilities of a dot occupying any position in the matrix, and these two probabilities were alternated from one strip to the next. The occurrence of a dot at any specific position in the matrix was independent of the occurrence of all other dots, so that each display was essentially a random sample from a set of independent binomial distributions. The program used to generate these displays could produce any number of such matrices, identical in all respects except that they were independent random samples under the same restrictions. The number of strips, the

two probability levels, and the size and density of the matrix were all easily controlled in the program.

Figure 2 illustrates the use of the computer to perform repeated nonrigid transformations on a pattern. In this program each point in the pattern undergoes a two-dimensional random walk from frame to frame. The walks for each point in the pattern are computed independently. With equal probability, each X coordinate may have added to it $+C$, $-C$, or 0, as may each Y coordinate. Between two frames, a point may thus move to one of nine positions, one of which is identical to its previous position. Each of these moves is equiprobable and, as in the previous example, the move for any dot is independent of all other dot moves. This process may be continued for as many walks as the program specifies, and the size of the constant is an easily controlled parameter. With enough walks an easily recognizable pattern may move into chaos, and when such a film strip is played backward, a recognition threshold obtained. This program has been used (White, 1960) starting with a pattern of randomly distributed dots in order to generate a film strip which could be used to give observers a controlled preexposure field of motion of a random direction and known average velocity.

Figure 3 illustrates a somewhat similar program, dubbed "Metamorphosis," designed to turn any pattern into any other in a given number of moves. Two patterns are stored in the computer as lists of X-Y coordinates and are paired according to some rule. A path is then computed from each point in pattern A to the corresponding point in pattern B and this path is divided into as many segments as the number of frames to be displayed. On each frame each point moves one step toward its destination in the second pattern. The second pattern can of course be a random set of coordinates within some given limits. In this case the first pattern heads straight for chaos instead of getting there aimlessly as in the random walk program.

The metamorphosis and random walk programs give some idea of the ways in which the computer can be used to perform

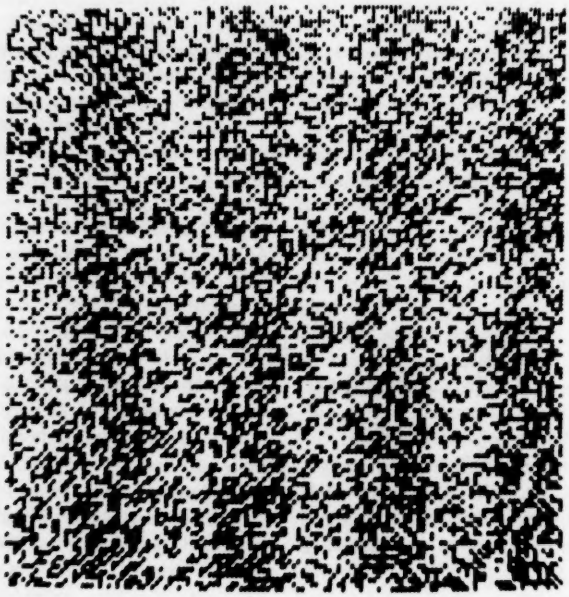


FIG. 1. Statistically Defined Bar Pattern (From Green, Wolf, and White, 1959)

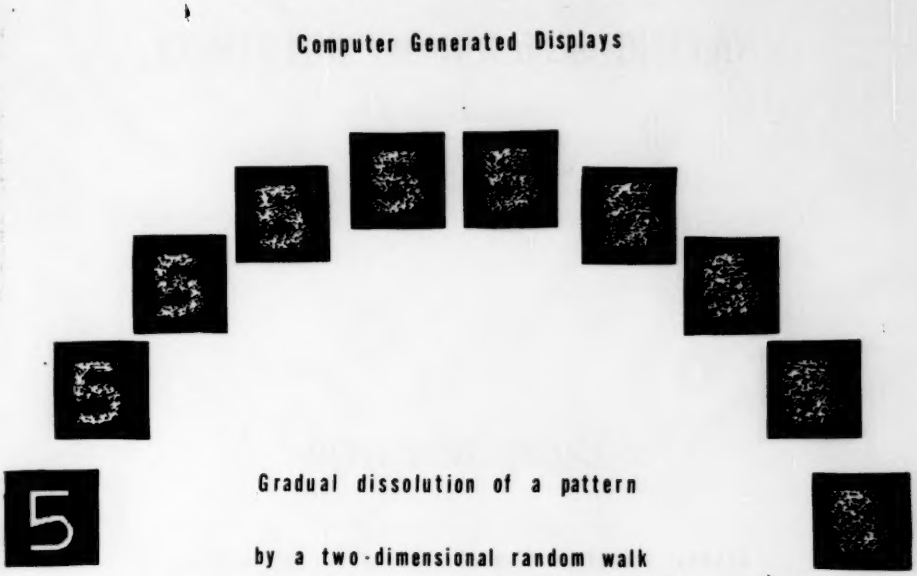


FIG. 2. Gradual Dissolution of a Pattern by a Two-dimensional Random Walk

unusual transformations on two dimensional patterns. One experiment has shown a most orderly relationship between the number of random walks an alphanumeric character has undergone and the probability of its correct identification.

Though the display generated by the computer is a two-dimensional pattern on a scope face, a pattern of any dimensionality can be represented in the computer's memory. The series of patterns shown in Figure 4 is taken from a program written by Bert Green to generate stimuli for the study of the kinetic depth effect (Green, 1958). Here the points in the pattern are represented by three numbers corresponding to the X, Y, and Z axes in three-dimensional space. The projection of these points on a two-dimensional surface is computed, and in this form the figure is displayed on the cathode ray tube and photographed. The program then causes the X, Y, and Z coordinates to be transformed in a fashion to correspond with a rigid rotation of the pattern in three-dimensional space. The two-dimensional projection of the rotated figure is then computed and displayed. In this manner the computer generates a movie frame by frame which, when viewed at conventional movie speed, is seen as the rotation of a three-dimensional figure in space. Green, in his program, had a number of parameters under program control: the number of points or lines in the pattern, the speed of rotation, the regularity of the pattern, and the direction of rotation.

Figure 5 is an extension of the same program to a four-dimensional hypercube. Here each corner is represented in the computer by four coordinates, this four-dimensional pattern is projected on a three-dimensional volume, and this in turn projected on a two-dimensional plane and displayed.

There are some limitations to this method of generating visual patterns which should be mentioned. A computer is fast, but getting a program written and debugged may not be, especially if, like me, you are an untrained programmer. Getting patterns into the computer is awkward without special equipment. All the X, Y coordinates have to be put on tape or cards in a form which the computer can digest. This can be tedious. There exist devices which can scan a photographic negative or print and which

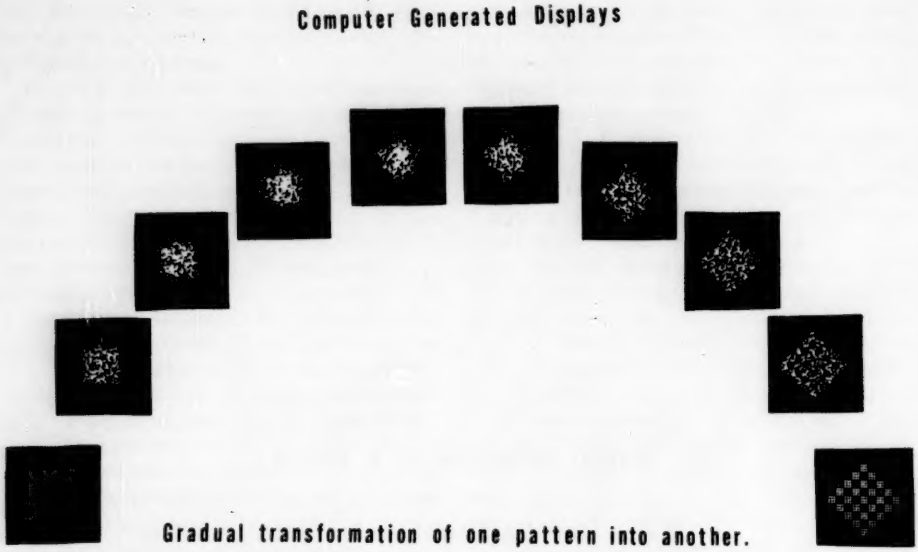


FIG. 3. Gradual Transformation of One Pattern into Another

store the output directly in the computer. For anyone dealing with patterns of great complexity, or with a large number of patterns, such a device would be a necessity. There is, also, as they say in this trade, an "output" problem. The display unit is on-line, and is a delicate thing requiring frequent adjustment. It is heartbreaking, but not unusual, to have a run on the computer, only to discover when your film has been developed that something went haywire with the cathode ray tube, or with the camera, so that your film is hash. For most purposes it would be preferable to have the computer put its output on tape and then use the tape to drive an off-line display unit, so that valuable computer time is not lost because of terminal equipment failure.

When one looks at the research done on perception, one is struck by the preponderance of studies in vision, although in recent years there has been a marked increase in work on tactual perception and on auditory perception, particularly in connection with speech recognition and intelligibility work. Most of the work in visual perception has been done with static patterns, largely because of the difficulty of manufacturing carefully controlled dynamic stimulus material. There are some notable exceptions, of course, like the work of Gibson who has done so much to exploit the motion picture as a device for presenting moving patterns, and the work of Philip and Fisichelli (1945) who studied apparent movement reversal using Lissajous figures generated on a cathode ray tube. The technique described here is an extension of this work. The moving picture camera has been placed before the tube face and the simple signal generators have been replaced by a computer which is capable of subjecting patterns to a wide variety of transformations. It may take several months to get a display program to work properly, but once debugged, and if you have been clever in writing it, it is very easy, at the flick of a switch, to change such display parameters as: size, speed of movement, texture gradient, slant, number of elements, severity of random perturbation. When patterns are produced by drawing or conventional photography, making a new set with altered parameters is often as

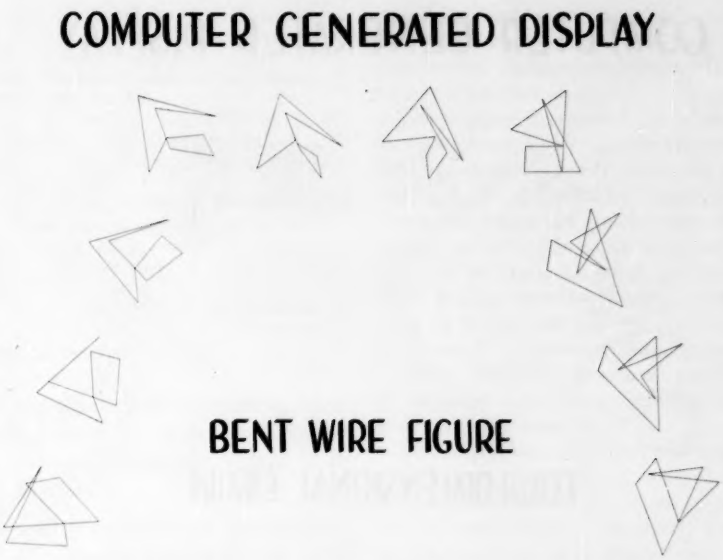


FIG. 4. Rigid Rotation of a Three-Dimensional "Bent Wire" Pattern

laborious and time consuming as producing the original.

It seems to be the case in perceptual research that if some aspect of a visual pattern is difficult to make, one tries to eliminate it from displays altogether. If we cannot create textures in a controlled fashion, we go to great lengths to eliminate texture from experiments on shape constancy, for example; or if we cannot control movement of the image on the retina, we resort to the bite board and fast exposures. The results of these kinds of drastic editing of the visual world are evident in most textbooks when such topics as space perception are discussed. In such discussions, it is easy to get the impression that the perception of depth is mainly a matter of accommodation, convergence, and retinal disparity, but that if you want to cheat, you may also move your head and look at the microstructure of the surfaces confronting you. Now that it is possible to manufacture stimulus material with such attributes as movement and texture under close control, there is no longer an excuse to keep our subjects in an impoverished perceptual environment in the laboratory.

There are two other ways in which computers may be used in this area of visual perception and pattern recognition which I think will become more prominent in the next few years. So far, I have talked only of the computer as a machine for generating film strips. It is quite possible to think of using the computer as a "live" stimulus pattern generator. In such use, the subject would sit before the visual display generated by the computer (which display might be static or moving). He would be provided with controls which in effect intervene in the generating program in such a way as to alter the display. Existing systems already provide observers with the capability of adding or eliminating dots from the display through the use of a probe containing a photoelectric element. It would be equally easy to provide controls for the rate of movement, size, or any other display parameter one wanted to put under subject control. In recognition studies, the subject could be provided with a choice of response buttons which, when pressed, would inform him of the correctness

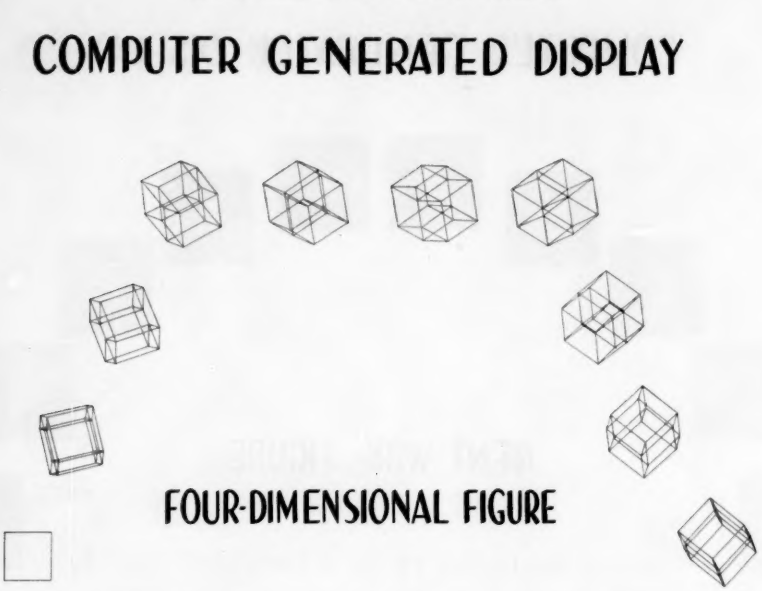


FIG. 5. Two-Dimensional Projections of the Rotation of a Rigid Four-Dimensional Figure

of his choice and cause the machine to generate the next display. Something very like this has been used for some time in the evaluation of guidance-control systems, and something very like this is certainly at the heart of the more elaborate teaching machines.

A great deal of effort is going into writing programs which will enable computers to recognize visual patterns. At our laboratory there are several such efforts. One, for example, has to do with computer recognition of spoken English (Forgie and Forgie, 1959), and another with computer recognition of hand printed letters (Doyle, 1960). Though each of these projects has started with some more or less explicit notions about how people recognize patterns, the objective has been to get an existing machine to do so in a reasonable length of time, not primarily to test a model of human pattern recognition. Such computer programs, however, are working pattern recognizers whose performance may be compared to that of human recognizers, and it is inevitable that such comparisons are going to give the programmer some ideas about improving the performance of his computer, and the psychologist some new ideas about experiments he wants to do with humans. Working models of a process as complicated as pattern recognition have not existed before, which is one reason why this area of psychology has been characterized by theories which are more speculative than rigorous.

It would be absurd to say that the computer is going to clear all this up. But it is not absurd to say that the computer provides a unique and flexible device for generating visual patterns for use in perceptual research, and, since it has already been used as a pattern recognizer, it can also serve as a way of sharpening and correcting theoretical models of human perceptual functioning.

REFERENCES

Doyle, W. Recognition of sloppy hand-printed characters. *Proceedings Western Joint Computer Conference*, 1960, 17, 133–143.

Forgie, J. W., & Forgie, C. D. Results obtained from a vowel recognition computer program. *J. Acoust. Soc. Amer.*, 1959, 31, 1480–1489.

Gibson, J. J. (Ed.) Motion Picture Testing and Research. Report No. 7, AAF Aviation

Psychology Research Reports. Washington: U. S. Government Printing Office.

Green, B. F., Jr. Some conditions for the occurrence of the kinetic depth effect. *Amer. J. Psychol.*, 1958, 13, 406.

Green, B. F., Jr., Wolf, A. K., & White, B. W. The detection of statistically defined patterns in a matrix of dots. *Amer. J. Psychol.*, 1959, 72, 503–520.

Philip, B. R., & Fisichelli, V. R. Effect of speed of rotation and complexity of pattern on the reversals of apparent movement in Lissajou figures. *Amer. J. Psychol.*, 1945, 58, 530–539.

White, Burton L. The role of visual experience in velocity perception. Doctoral dissertation, Brandeis University, 1960.

A Computer-Controlled Teaching Machine, *Harry F. Silberman, System Development Corporation, Santa Monica, California*

This is a progress report on a project to develop a computer-controlled teaching machine to be used as a research tool for the study of different teaching procedures. The plan for this machine was described at the American Psychological Association meetings last year (Silberman, 1960). Since then, several computer-controlled teaching machines have been announced, including one being used behind the iron curtain to teach foreign languages (Central Intelligence Agency, 1960; McCarthy, 1959; Kopstein, 1960). Our machine has been in operation for several months now. It is our hope that, after a considerable amount of research with this machine, a sufficient number of answers will be available to permit the development of a simple special-purpose teaching device for applied use.

Briefly, the machine consists of three elements: a Bendix G-15 computer, a large random access slide projector, and an electric typewriter. The student reads instructional items projected from 35 millimeter slides by the random access projector. He then enters his multiple choice response on the typewriter. His answers are recorded and evaluated in the computer by comparison with answers that have been pre-stored on punched paper tape. The paper tape contains information about each item in the projector, and is an input to an internally stored slide selection program. The student is supplied knowledge of results either by a printout on the typewriter or by the presentation of another slide.

The computer selects a variable sequence of slides from the projector in response to different aspects of each student's behavior. For example, if the student makes too many errors, or takes too much time on certain topics, or, in reply to an item asking him how well he thinks he is doing, indicates that he is having trouble, he will then receive extra remedial items. If the student is making no errors, responding rapidly, and feeling confident about his progress, then items will be skipped permitting him to cover the material more rapidly.

The items are organized into topics, each topic being covered at several difficulty levels. Items covering a topic at any given level cover the topic completely. If a student is doing exceptionally well at one level of a topic, he may be branched to a more advanced level of the topic. If he is not doing well, he will be branched down to a lower level of items which use a different approach to the topic. If he does poorly on this level, he will be shunted to still lower levels prior to moving on to the next topic.

The machine is responsive to the individual student in that slower students are branched to lower levels while the upper levels constitute an expressway for the fast student. This differs from the popular approach which requires all students to traverse all items in fixed sequence irrespective of individual differences (Skinner, 1954).

Some of the multiple choice items contain distractors which indicate a common type of misunderstanding. If one of these distractors is chosen, the student is immediately branched to remedial items which are designed to clarify that particular point. At the end of each topic, or at the end of the lesson, the slides with which the student had trouble may be repeated for review purposes. The student may also change his answer to a question prior to inserting it into the computer, or he may cause the machine to back up, one slide at a time, to review some material before entering his answer.

The basis for the design of a teaching

machine in which the sequence of instructional items is tailored to the needs of each student is the notion that there is no single best item sequence nor one best method of presenting the materials. The importance of individual differences lies in the interactions of the students with method and task variables. As research with this machine reveals dimensions of individual differences which interact with method and task variables, the machine can be programmed to treat students differentially with respect to such dimensions. This tracking of the student's performance is expected to make a substantial contribution to student learning.

Data have been gathered on the effectiveness of some of these program features from experiments in our laboratory this summer. Four hundred multiple choice items on logic were prepared and given a preliminary trial in a number of schools. Items were revised and put onto thirty-five millimeter slides for use in the random access projector. The results support the view that the intelligence of the student is positively related to the effectiveness of the machine as a teaching device when each student is given only one trial on the material to be learned.

The data that have been collected also emphasize the importance of careful ordering of the set of instructional items from which the computer will make its selection. Although item sequencing still seems to be an art, some consistencies show up on error analyses which suggest the feasibility of more systematic approaches to sequencing of items. For example, in teaching a logic lesson, the student was presented with various arguments and asked to determine their validity. The arguments consisted of premises and conclusions which were true in some items and false in other items. Students had been instructed previously that the validity of arguments depends on the form of the argument rather than on the truth or falsity of the premises or conclusions. Yet items which present a familiar true statement as the conclusion of an invalid argument were usually missed. Students tended to accept the argument as valid even though preceding items instructed them to attend to the form of the argument rather than to the truth or falsity of the conclusion or premises. Students would not reject invalid arguments which appealed to their common sense. When these same invalid arguments were cast in symbolic form very few errors were made. This also applies to the case where students rejected valid arguments which appeared faulty. The sequence of items should probably be ordered so as to begin with the symbolic forms of the argument and end with the verbal forms of the argument.

Another experiment in our laboratory this summer, which did not use the machine, demonstrated a difference in favor of a branching capability which allowed subjects to review earlier material, compared with a fixed sequence of items without the freedom to branch. Subsequent studies will utilize different branching methods, a revised set of teaching items, and additional program features such as variable feedback messages, machine responsiveness to the student's response latency, and a capability for student questions and self-evaluation.

Perhaps a few words should be said about the limitations and advantages of using a computer-controlled teaching machine as a research tool. To begin with there is of course the question of expense. The monetary cost of a computer laboratory is high. Furthermore, the equipment will present maintenance problems and the researcher soon becomes dependent upon technical personnel. It appears simpler to plan an elaborate computer facility than to develop it. There seems to be some law which makes computer systems less ambitious than their designers initially envisioned. The size of the computer is a limiting factor. If the storage capacity of the computer is small, the machine may allow only one subject to be run at a time. Finally, there is the possibility that the equipment may dictate the research which is attempted. It is necessary to guard against selecting problems for which the equipment is well suited but which have only slight relevance to the teaching-learning process. It is also inefficient to use the computer laboratory to test hypotheses which can be readily attacked with much simpler apparatus. The computer does not diminish

the need for creating imaginative research hypotheses to attack the traditional learning problems which are still with us in teaching machine research.

On the other hand, one of the reasons for using a computer was the difficulty of doing the same job manually. The computer is a general purpose instrument which processes any information that can be coded for it. This very coding process forces operational specification of the teaching process in considerable detail.

Other advantages of the computer as a research tool include the following:

1. The capacity of the computer to be programmed to modify the sequence of instructional items during the tutoring session in response to student performance. This permits adaptation to individual differences in a much more complex manner than could be accomplished manually.

2. The rapidity of computer operation in processing student response data. This is particularly important in systems where the computer is multiplexed to many student input stations.

3. The computer provides improved control over experimental conditions.

4. The computer can be used for reduction and analysis of experimental data during and after the training session.

5. The flexibility to change the operation of the equipment by program modification. This allows both existing and proposed teaching machine features to be readily incorporated in the machine.

6. The capability of obtaining complete records of the internal operation of the machine during the teaching session.

Perhaps one of the more important functions of the computerized laboratory for educational methods may be to test hypotheses prior to extensive field tests in the school setting.

While at present this is only a research tool, we may contemplate powerful miniaturized computers of the future processing data obtained from large numbers of students with a consequent increase in productivity per man-hour of teaching effort.

REFERENCES

Central Intelligence Agency, Scientific Information Reports, June 3, 1960, Pp. 132. Distributed by Office of Technical Services, Department of Commerce.

Kopstein, F. F. The scientific teaching of scientific material. Paper presented at the *National Aeronautical Electronics Conference*, Dayton, May 2, 1960.

McCarthy, J. Personal communication. December 28, 1959.

Silberman, H. F. A computer as an experimental laboratory machine for research on automated teaching procedures. *Behav. Sci.*, 1960, 5, 175–176.

Skinner, B. F. The science of learning and the art of teaching. *Harv. Educ. Rev.*, 1954, 24, 86–97.

Computer Simulation of Interaction in Small Groups, *A. Paul Hare, Haverford College*

For the past few years, the half dozen or so persons who have worked with Freed Bales on various research projects at the Laboratory of Social Relations at Harvard have had an interest in computer simulation of interaction in small groups. Bales has been following the development of the Newell-Shaw-Simon IPL-V Programming Language (Green, 1960) and has suggested that the goal of naturalistic prediction of behavior in small groups may only be possible through the use of computer simulation (Bales, 1959). To this end, Bales has been developing a comprehensive outline for a program for simulation that he has called the "Interaction Simulator." The goal of the Interaction Simulator is to reproduce the actual content and process of small group discussion once the personality characteristics of the members and the discussion topic have been specified. That is, the output of the computer will be a series of statements each indicating who is speaking, to whom he is making his comment, and the task and social-emotional content of the comment. A series of statements representing 40 minutes of interaction could then be compared with the interaction of real people holding a discussion on the same topic. This is a long-range goal, a goal that has only appeared to be attainable through the use of high-speed computers.

Where are we now in relation to this goal? Bales's master program, the Interaction Simulator, is well underway. Art Couch, through his Ph.D. thesis, has made a major contribution on the side of content by doing a series of factor analyses to discover the personality and interaction factors that must be considered in the simplest model of social interaction (Couch, 1960). On the side of method, Couch has written computer programs which cover all the data-processing steps necessary to monitor and analyze the output from a five-man discussion group.

Over the past two years while we were waiting for IPL-V to be debugged, I decided to try out some of the smaller pieces of the puzzle on the IBM 650. Although it lacks the storage capacity of the large machines, the 650 nevertheless performs all the basic operations and has the added advantage that one can deal directly with the machine as programmer and operator. My programs are written in machine language. I decided to begin writing programs for the task activities of individuals, then of groups, and then to begin adding programs to simulate the social-emotional behavior which goes on concurrently with the task behavior. The group to be simulated is a five-man group of college undergraduates with a discussion task, the "standard" laboratory group.

TASK ASPECTS OF INTERACTION

One task which has been developed for use with five-man laboratory groups requires the subjects to predict the responses of an unknown student on the Bales-Couch Value Profile (Bales & Couch, 1960). Group members are given the answers to five items from the value profile and then asked to predict the responses of an unknown student to ten or more additional items. The possible responses are: strongly disagree, disagree, slightly disagree, slightly agree, agree, and strongly agree. Among the five items might be: "Each one should get what he needs—the things we have belong to all of us" which the unknown student answered "Strongly agree"; and "You have to respect authority, and when you stop respecting authority, your situation isn't worth much" which the unknown student answered "Disagree." After the five group members discuss the answers to the five initial items, they are given a new set of items and asked to predict the responses. After each prediction, the actual answer to the item is revealed. An item to predict might be "Obedience and respect for authority are the most important virtues children should learn." Those of you who are familiar with the "F" Scale will recognize this as an item representing the value "Authoritarianism." Knowing this, you would probably predict that the unknown student answered this item "Disagree," as he did the "Authoritarian" item in the example above. Actually, he answered "Strongly disagree" and you would have an absolute discrepancy of one point on the first item.

The particular unknown student in the example above was a graduate student in clinical psychology who was chosen because he held strong and rather inconsistent opinions. If you are a person who is familiar with the Value Profile, or if you are a properly programmed computer, you know that the four principal value factors are: (I) Acceptance of authority, (II) Need-determined expression vs. value-determined restraint, (III) Equalitarianism, and (IV) Individualism, and that the test items represent these factors or combinations of them. With this knowledge you would be able to do fairly well in predicting responses. If you are a Harvard undergraduate, you will probably not be able to do as well but you will find it an interesting task.

The task of simulation began with an attempt to reproduce the process each group member would use in reaching his own individual decision on an item before he announced his prediction to the other group members. This program could also be used to simulate his performance as an individual on the same task since for some groups we have used individual performance on this task as a premeasure of the subject's task ability. The question is, by what process does the subject make his predictions? One possibility is that the subject looks for the previous item most similar in content to the new item and predicts that the unknown student will answer in a con-

sistent fashion. In the machine, the content of an item can be represented by the four factor loadings on the major value factors. Phil Stone of this group has written a program which, by correlating the four factor loadings for a new item with the factor loadings of all previous items in turn, identifies the item having the highest correlation with the new item. This approach is fairly successful but it does not take into account response tendency, since it considers only one item at a time, and it slows down rapidly as the number of items to be considered is increased.

Another possibility for which I have written a 650 program is that the subject first identifies a new item as belonging to one of a limited number of categories which include the four principal value factors and their major combinations, and then predicts that the unknown student will answer the new item as he has answered similar items "on the average" in the past. (The average of the previous items was first calculated by taking the median. This was later changed to the mean when I found that the mean gave a more accurate representation of real subjects.) Since the machine—as well as the group—is given initial items which cover only the four major factors on the Value Profile, new items representing combinations of these factors are not contained in its memory. In such a case, for example an item which is highly loaded on both (I), Authoritarianism, and (III), Equalitarianism, the machine is programmed to take the mean of all previous I and III responses. Since there is no zero point in the response form for each item the machine is not allowed to guess "zero." If the mean of all previous responses happens to be zero, the machine looks at the unknown student's response set to see whether he has a tendency to agree or disagree and then predicts accordingly. With a "predictable" unknown student, the machine makes an average error of about .71 over a typical run of a dozen predictions. The program takes about 400 words of storage. The input is one card for each item giving the four factor loadings for that item and the true answer. The output is a card giving the category the item is guessed to be, the guessed answer, the discrepancy, the average discrepancy for all items up to that point, and the response set. At present, this program does not allow for individual differences in problem solving, although we have some ideas for possibilities to try.

A second piece of the puzzle is a program for the IBM 650 to simulate the decision reached by the group on the prediction for each item. This program, written with Elizabeth Borland, begins with the five individual predictions as inputs. The program assumes that the group members typically reach a decision that represents an average of the individual members' judgments. As before, if the mean of the individual judgments is zero, the program checks the response set. If the response set is also zero, the machine determines whether the majority of the group members thought the unknown student would answer on the agree side or the disagree side of the scale. This is a smaller program than the first, about 100 words. The input consists of a card with the five individual predictions and the true answer for each item. The output consists of a card with the true answer, the guess, the average discrepancy, the response set, and a note if the "majority opinion" loop was used.

As a test of the effectiveness of these two "simulations," the predictions generated by the machine program were compared with actual predictions of individuals and decisions of groups of 20 laboratory groups of five Harvard undergraduates. These data were collected by Lindsey Churchill (Churchill, 1961). The group members met for one 70-minute discussion. Most of the members were unknown to each other before the experiment began. Each member was given a card with one item from the Value Profile and the answer given by the unknown student (the clinical psychologist). After a short discussion of about 10 minutes, during which the members pooled their information about the unknown subject, they predicted his responses to an average of about 10 items, depending upon the speed of the group. Before each group decision was

made, each subject recorded privately his own prediction.

The average discrepancy between the guess and the true answer for the 100 experimental subjects was 2.40. The average discrepancy for the computer making the same predictions was 2.00. Although better than the average subject, this is still not very good. Recall that the unknown student was chosen because he was thought to hold strong opinions and to be somewhat unpredictable. This hypothesis is confirmed. He was indeed unpredictable. The average discrepancy score for the 20 men who were the most accurate in their groups was 1.85, slightly better than the machine. Apparently the machine still has something to learn.

A comparison of the group decision on each item with the true answer indicates an average discrepancy of 2.38, about the same as that for the average individual. This means that the results of group discussion do not differ significantly from the accuracy obtained by simply pooling the individual guesses. Since in every group the subject who was most accurate was not the one whose answers were closest to the group decision (in two groups the most accurate man was tied with another), it is evident that the groups were not making full use of their most insightful members. The groups which were the most accurate were the ones which contained, on the average, the most accurate individuals. These groups also completed the largest number of predictions in the allotted time. Similar results of comparisons of groups and individuals in matters of judgement have been reported in previous research (Kaplan, Skogstad, & Girshick, 1950).

The members of the 20 groups made a total of 180 group predictions. Of these, 112, or about two-thirds, were reproduced exactly by machine simulation that took the mean of the five individual predictions. In those cases where the mean was different from the actual group decision, the machine was closer to the correct answer about half of the time. During the analysis of the data, it appeared that a program which took the majority opinion when there was a clear majority might be closer to the actual group decision than one simply taking the mean. If the majority opinion had been considered first, the simulation of the group prediction would have reproduced the actual group prediction in an additional nine instances (or about 5 per cent). However, the "majority" routine would have been less accurate than a routine which always took the mean in predicting the true answer by almost the same amount, in eight instances. One concludes that a routine that took the majority opinion first and took the mean of the individual guesses when there was no clear majority would have improved the simulation only slightly.

Since one third of the group predictions are not reproduced by the present simulation, it would be desirable to find some other source of unaccounted-for variance. Is it possible that these groups have leaders who are influencing group opinion? In nine of the groups, we would have done better to use the guesses of the subject who was closest to the group prediction than to use the mean of all the individual guesses. If we could identify these men before the experiment, we might be able to improve our predictions. Unfortunately, with only nine such persons in our sample it is difficult to draw any definite conclusions. Although seating position and pretested personality measures have been found to be related to amount of talk and amount of influence in other experiments, these nine subjects do not show distinct preferences in seating or distinct personality characteristics for the dimensions of anxiety, warmth, and aggression. There is one possible exception in that five of the nine rank lowest in their groups on the dimension of aggression. However, this does not fit our impression of a very influential type. In any event, this is only the beginning; we will continue to try to improve the simulation.

SOCIAL-EMOTIONAL ASPECTS

So far we have talked about the task side of individual and group performance. What about the social-emotional side of interaction? The final simulation should consider not only what is said but also how it is said.

We have a beginning in a program written by Sherwin Feinhandler and me to score the Conditional Personality Test, a new type of test developed by Paul Breer (1960). The test has been used only once with subjects in Breer's thesis research, some 20 groups of five persons each. The groups contained both men and women undergraduates. The test consists of four paragraphs describing persons with different personality characteristics: one dominant, one submissive, one positive, and one negative. Each paragraph is repeated twice, once with an indication that the person described is male and again with an indication that the person is female. The subjects taking the test are asked to indicate how they would behave if such a person were in a discussion group with them; specifically, how much they would like him, how quiet or forceful they would be, and how critical or warm they would be. The test is further complicated by asking the respondents what difference it would make if the other person were different with respect to age, intelligence, social class, race, looks, and values.

There are two ways to use the test results. One is to average the responses for all four personality types for both sexes to determine how a subject would behave toward others in general. The other way to use the test is to select only those responses which refer to the characteristics of the other subjects who are actually in the group and then compare the test responses with the actual behavior. This latter task required a fairly elaborate program of some 800 words to do the actual scoring with an additional input-output program of 200 words written by Couch. Unfortunately, when the program was tested with Breer's data, the final conditional score for each subject for dominance and warmth did not differ by more than a fraction of a point from the unconditional score. We shall have to try again.

Although this program was written primarily to score a specific personality test, it also provides a piece of the total simulation program. It allows us to simulate the initial perceptual part of the social act as the new group member looks around at the other members and decides on the basis of his past experience how he will act toward each one and how as a result he will act in general.

And so we proceed. Some of us have left Harvard, but Bales and Couch remain and others will join them. In my experience, the computer is a hard master since it forces one to be specific about the variables in interpersonal behavior and the exact relations between them. But its special value lies in its ability to make predictions by considering in a systematic way far more variables and relationships than an individual could hope to handle in the short run. Thus it becomes possible to make real predictions in real time about social interaction in small groups.

REFERENCES

Bales, R. F. Small-group theory and research. In R. K. Merton, L. Broom, & L. S. Cottrell, Jr. (Eds.), *Sociology today: Problems and prospects.* New York: Basic Books, 1959, Pp. 293–305.

Bales, R. F., & Couch, A. S. The value profile: A factor analytic study of value statements. Harvard University: Laboratory of Social Relations, (ditto), 1960.

Breer, P. E. Predicting interpersonal behavior from personality and role. Unpublished doctoral dissertation, Harvard University, 1960.

Churchill, L. C., Jr. Aggression in a small group setting. Unpublished doctoral dissertation, Harvard University, 1961.

Couch, A. S. Psychological determinants of interpersonal behavior. Unpublished doctoral dissertation, Harvard University, 1960.

Green, B. F., Jr. IPL-V: The Newell-Shaw-Simon programming language. *Behav. Sci.*, 1960, 5, 94–98.

Kaplan, A., Skogstad, A. L., & Girshick, M. A. The prediction of social and technological events. *Publ. Opin. Quart.*, 1950, 14, 93–110.

Empirical Inputs in a Voting Model, *W. N. McPhee, Bureau of Applied Social Research, Columbia University*

Abstract.[1] A computer model of the formation of opinion in response to election campaign appeals was used in 1960 to conduct experiments. The model accepts as initial input a miniature electorate who are

[1] The full text of this paper by Dr. McPhee appears in the *Public Opinion Quarterly*, July, 1961.

respondents in a sample survey and studies changes in opinion as a result of (a) internal processes of the model such as discussion and learning, and (b) external inputs of new stimuli representing subsequent campaign appeals. The paper deals with the latter.

In the Wisconsin presidential primary election in 1960, Senator Humphrey stressed the "farm" issue as an argument in his behalf. Interviews showed that some Wisconsin people thought this was a strong argument, some fair, some poor. An appeal thus is a distribution of appeals. The model accepts such distributions, obtained from surveys, and converts these into "effects" which are monotonic functions of a basic disposition or probability of voting for Humphrey. The *form* of this function changes in a favorable or unfavorable way as a result of the perceived cogency of the appeal. The "effect" of an appeal or argument is thus defined as the change in form of the function which relates responses to the basic dispositions of the electorate. Results from the Wisconsin primary are used to illustrate the model.

Comments on the Use of Computers in Psychological Research, *Nancy S. Anderson, University of Maryland*

For the past several years, there have been at least one, if not more, symposia at the American Psychological Association meetings whose topics of discussion have centered around the use of computers by psychologists. While employed in a computer industry, I used to hear and read a number of papers on the uses of computers for a multitude of business applications as well as scientific applications. I have noticed a trend over the past few years: the papers, symposia, and articles *now* appear in specific content areas. That is, the computer has become accepted as a tool in scientific research, and new applications are described only in connection with the specific content area of the discipline. For example, in psychology, there are papers and articles in the content areas of Teaching Machines and Information Processing Models of human behavior. These papers describe the use of computers for their particular research purposes. Thus, we are talking less and less about the computer itself and more and more about the research we are doing while using this computer only as a tool. It is reasonable at this time to ask why we use the computer as a tool.

Clearly, in the area of statistical and data analysis, the computer is a useful tool for the psychologist whose desk calculator has now been referred to as a "relic" of the past. It has even been stated that the 650 is a slow, slow, . . . computer. Perhaps in ten years, someone will tell how slow and passé the fast computers of today are. Nevertheless, if any tool does its job, it is a useful tool; if you want a faster one, find one. If, however, you can't afford a better or faster one, you can still continue to do research with what you have. Certainly, not everyone drives a Cadillac, and there is even a trend towards the smaller less high-powered cars today. There is a danger in the high speed computer—it gives you an enormous listing of all of your calculations, and now you must immediately spend time deciphering them before the data go "cold" and you forget what you asked the computer to do for you.

Now, let us look at the area of computer use that we are concerned with in this symposium. Here, we have been shown a wide representative class of problems which use this tool as (1) a pattern generator, (2) a control program for the teaching machine, (3) a simulator of small group interaction behavior, and (4) as a model for voting behavior. In each of these cases, I will discuss the nature of the use of the computer as a tool and attempt to point out what has been done to advance the state of the art that could not have been done without the computer.

In the field of vision and perception, one of the problems facing any researcher is how to generate stimulus materials. This is especially true if the stimulus materials are complex patterns. The work reported by Dr. White indicates the wide range of patterns that can now be generated by the computer through a cathode-ray-tube output and a motion picture camera. The advantages he

pointed out are very clear, such as, the accurate control over parameters of the stimulus patterns, and the ability to change these parameters easily with a well-written computer program. It is true that much of our pattern perception research has been restricted simply because we could neither control nor generate certain parameters of complex patterns in any practical way. Now that we can produce parameters, we can learn more about perception by studying it with a more realistic representation of the visual and auditory world.

I hope that as Dr. White and his colleagues pursue their research they can utilize the computer as a "teaching machine" in their perceptual experiments. Then, the experiments can be conducted to provide rich environments and the means for the subject to respond to these environments. Some psychophysical experiments have been "automated" in this manner by Dr. Tanner and his associates at Michigan and by Dr. Boyton at the University of Rochester.

Dr. Silberman has described a computer-controlled teaching machine, which consists of the Bendix G-15 computer, a slide projector for presenting stimulus materials, and a typewriter. One might ask why one needs a more complex teaching machine than those used by Dr. Skinner and others. Dr. Silberman has stated quite clearly the advantages and disadvantages of this teaching machine system. Few will disagree with his points, but perhaps a few more can be added, and others re-emphasized.

One of the most talked about advantages of the teaching machine approach is that the whole process of learning is adapted to individual differences of the student. It is clear that a more complex device like a computer can more easily accomplish a complete or complex change in the teaching program to suit individual needs of the student. If one wishes to alter the presentation of material by using a criterion that is more complex than merely the rightness or wrongness of a particular item, the tool to control the teaching machine will have to be more complex. The computer can accomplish this by a well-written computer program which modifies the computer program which in turn controls the sequencing and nature of the material that is presented to the student.

One of the major bottlenecks in the development and use of teaching machines today is the preparation of the teaching programs. First of all, a program takes time to write and then must be modified and tested and revised until it meets the writer's criteria of a "good program." This task is so time consuming that it would be very advantageous for the computer to aid in the process. If the computer were used to do complete item analyses on all students' responses, then at least the task of revising the items and their sequencing could be accomplished more easily. Furthermore, there must be some way to use the computer to aid in generating teaching programs, that is, the actual items and/or the sequencing of the items. Once these procedures can be worked out to help in program building, then the computer can be used in real-time interaction with the student as a data processor and record keeper for the learning process of each and every student. This discussion sounds a little like automating the teaching machine business, but this will probably never happen. However, the complex task of building teaching programs could use some kind of help from our tool, the computer.

In the area of simulation of behavior, Dr. Hare has given a description of his experiences in the use of a computer to simulate behavior in small groups. In addition to the work that Dr. Hare and his colleagues have done in simulation of behavior, most of us are familiar with similar efforts such as those of Rochester, Holland, Haibt, and Duda (1956) and their simulation of Hebb's model of a neuron net; Rosenblatt (1958) and his work with the "Perceptron"; and Newell, Shaw, and Simon (1958) and their work in the general area of simulation of problem solving behavior. Dr. Hare's work is most similar to that of Newell et al. (1958) in that the simulation of behavior on a computer is conducted on a descriptive as well as a mathematical level. The models for problem solving and small group interaction use an information processing language which is not restricted to mathematical

operations only, but can perform logical operations as well. In his use of the computer, Dr. Hare describes group behavior and uses the computer to perform numerical calculations on a large number of variables that might not have been considered in the model had he been unable to use the tool.

In addition, the computer has forced him and others who use the computer for simulation to be accurate and precise about the variables and their relationships. The computer is a hard task master. If one chooses a model for behavior and writes mathematical equations, one is very specific. I think that one is forced to be just as specific and rigorous when programming a digital computer to simulate human behavior. Thus, the computer contributes to the process of model building by forcing the theorist to specify precisely the parameters of behavior. This is a healthy trend, and may help us become clearer in some of our descriptions of behavior.

There is at least one problem that should be mentioned with respect to this idea of the computer as a hard task master. If the computer simulates behavior, does it necessarily mean that this is exactly how the person behaves? It is not necessary to pursue this point too far in criticizing the usefulness of model building, since Dr. Chapanis in his Presidential Address to Division 21 of the American Psychological Association stated clearly both the issues and problems of determining what is a "model" of behavior. It is important to emphasize one point. The language used to simulate the model on the computer may not be the language needed to describe the behavior. The languages used to program computers, even FORTRAN and IPL, are designed to carry out the computer operation efficiently. It may be that these languages, in addition to the computer's logical operations and controlling functions, force us into describing behavior in an artificial manner. Thus, though the computer forces us to be precise, the manner, and not the extent to which it does this, may be inappropriate or not necessary for a model of behavior.

Dr. Hare hopes to overcome part of this problem by testing specific hypotheses for his model and comparing them to actual human behavior. But what is at fault if the prediction is not good—the model, the way the computer forces us to simulate the behavior, or the measure of the behavior? This question cannot be answered without further research, but we must be aware that we are "simulating" and attempting to predict only a part of behavior, and our simulation is restricted by the specific characteristics of the "tool" and its capabilities.

Many of the ideas discussed above apply also to the simulation and model building described by Dr. McPhee. Dr. McPhee has discussed simulation of the behavior of the American voting public. In his model of voting behavior, certain functions of the parameters are established which describe the behavior. He uses survey data about issues or new happenings in the national picture to cycle the model ahead in time. In this way, he grinds out predictions for the future simply by using the computer to help him make changes in the parameters of the model. Not all parameters need to be changed, since he incorporates an internal bias which sets a frame of reference. In certain cases, a very small change in the national news picture or in survey data could result in a change in the parameter, depending on the bias or "predisposition" of the voting model. In his model, he attempts to account for the dynamic changes in voters' preferences. The complexity of the model is dependent on the number of parameters necessary *and* on how much information from surveys and the news of the world will be fed to the model. This approach has intuitive appeal since voters' behavior is accounted for by more than the Yes, No, and Don't Know approach.

Dr. McPhee needs a computer to help process all the data from surveys and national news and to convert the data into a form that can be used as an input to the voting model. The complexity of the model demands a large amount of computation in order to cycle the model for future conditional events. Now, the function of the model is not merely to predict the vote

count. He is actually predicting the manner in which the dynamics of the whole voting situation influence voting behavior. Looking at the flow chart of his model, it is clear that the use of the computer is dictated by the complexity of the model and the large amount of data processing and calculation that must be done. To carry out the dynamic cycling of his model might actually take years without the use of the computer. In this sense, his use of the computer as a tool is similar to Dr. Hare's. Without the computer, he could never account for this many parameters which influence the behavior and finish the computations for the model within any reasonable and practical time period.

There are other uses of the computer as a research tool that have not been covered here which are of general interest to psychologists, and as such, deserve mention. In the field of research on the training of executives, the American Management Association has developed a Management Game which has been used rather widely by some companies. More specifically, in the training situation, the human decision maker chooses a course of action for business decisions, and the computer calculates the outcomes of his choices. In order to carry out the computations, the computer has to have knowledge about the competitors' operations, past decisions of the business, and other parameters of the economic situation. The decision maker, in a sense, "operates" his business over time (usually set up in quarter periods), and throughout these operations gains some insight into the effects of his particular decisions on the economic status, productivity, and capacities of his business.

In another area of gaming, that of War Games, certain military tactical and strategic decisions for both adversaries are made and this information is fed into a computer. The computer calculates the outcomes of such decisions and can give the results of outcomes with certain theoretical assumptions for an unlimited variety of battle plans under various conditions. The computer can be used to evaluate the quality of the specific military decisions by choosing the criteria for winning. The entire procedure of assessment can also be used to evaluate future weapons against known and assumed defenses. Again, the results give the military decision maker some insight into the effects of his operational decisions. Why the computer is used for military war gaming tasks is quite clear. From a cost point of view, one cannot afford to carry out mock battles with missiles and other weapons systems. Thus, the use of a computer for war games allows the military personnel to carry out war operations under a variety of conditions and at a reasonable cost.

In my opinion, the reason we use a computer for a pattern generator, to control a teaching machine, and to simulate various models, is the same reason we use one for our statistical data reduction; as, for example, in pattern generation, the need for a device which helps us generate complex stimulus materials with a high degree of control over the many parameters of the stimuli. The capabilities of the computer and the associated input-output equipment make it well suited for the task Dr. White and his colleagues have at hand. Certainly, the computer has features that Dr. White doesn't use, but can one devise another way of generating his complex patterns and displays without building a special purpose computer whose cost would be prohibitive?

In the area of teaching machines, one wants to adapt and change the program to fit the student according to his performance throughout the lesson. This change of program should probably be done while the student is proceeding through the program. The computer can accomplish this for each student without physically changing the teaching program, but modifying it through its progam. Certainly, a special device could be built to accomplish the same result, but the computer can also perform data analysis while the student is learning. What the computer adds over a special device is its versatility in the capacity to act both as a control and data reduction device.

For the model builder in psychology today, there are such a large number of variables or parameters which are used in defining behavior, that the development of a

model for all possible values requires so much computation that it would be impractical without the use of the computer. Thus, we use the computer because the task we set out to accomplish is more complex than it was previously; and to finish the work within any reasonable time period, one needs the aid of a tool such as the computer. In each case that has been discussed at this symposium, it is clear that the computer was chosen as the tool that was most likely to help the researcher finish his work.

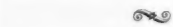
In conclusion, the computer is a tool; it has the same status, although more expensive, as that of the memory drum or desk calculator. It will never solve all of our problems in psychology, and it should not be used unless it will aid the researcher in his particular work. One should not use the computer for generating all stimuli for pattern recognition research because some patterns can be built more simply and more cheaply than by using a computer. Also, at present, not all teaching machine research programs will use the computer as a controller and data analyzer. Furthermore, not all model building in psychology will be done on a computer. One should consider the factors involved in his research task and let his own needs determine whether he uses a computer. If he chooses a computer, it should be used wisely, since it is time consuming and expensive. It is how much more one learns about behavior that can offset the cost of using a computer. One should not use a computer because it is in vogue, but because it is a tool that aids in the search for knowledge.

REFERENCES

Newell, A., Shaw, J. C., and Simon, H. A. Elements of a theory of human problem solving. *Psychol. Rev.*, 1958, 65, 151–166.

Rochester, N., Holland, J. H., Haibt, L. H., and Duda, W. L. Tests on a cell assembly theory of the action of the brain, using a large digital computer. *IRE Trans. Information Theory*, 1956, IT-2, 80–93.

Rosenblatt, F. The Perceptron; A probabilistic model for information storage and organization in the brain. *Psychol. Rev.*, 1958, 65, 386–408.



"The idea that the final task of science is the tracing of relationships of cause and effect has only an ill-defined meaning, in the light of our present concepts of the universe. But however meaningful such a view might be, this task in any case would not be the overriding challenge. That primary challenge, for science, is conditioned instead by its delight in the search for grand recurrences in nature. Such a search for regularities must ever, in the final analysis, resolve to the process of prediction."

CARYL P. HASKINS. *The Report of the President of Carnegie Institution of Washington*. 1959–1960

ABSTRACTS OF CURRENT LITERATURE WITH INTERDISCIPLINARY IMPLICATIONS

Prepared by Members of the Staff of the Mental Health Research Institute

SCHEIN, EDGAR H. **Interpersonal communication, group solidarity, and social influence.** *Sociometry*, 1960, 23, 148–161.

This paper grew out of the author's studies on Chinese Communist techniques of controlling civilian and military prisoners. It is concerned mainly with the importance of maintaining social relationships, roles, and self images in a group, for the stability of the group and of the individual personality.

The major hypothesis is that when interpersonal cues that confirm social norms, the individual's beliefs, self image, and basic values are absent, an individual becomes alienated socially. This alienation, this lack of reinforcement of the self image, makes him susceptible to influence because the forces that usually operate to make him resist such influences are removed. Social alienation is akin to a psychological state of disequilibrium.

One of the basic tactics used by the Chinese Communists on the United Nations prisoners of war was to destroy the authority structure of the group. They did this by isolating the leaders and key personnel from the remainder of the group. Low ranking enlisted men were chosen as squad leaders. The high ranking officers were threatened with punishment of their group if they did not co-operate with the Chinese with respect to different propaganda. When compromises were worked out by the high ranking officers to increase the chance of survival of their group, their actions appeared to the lower ranking officers and enlisted men as collaboration.

Another method used to interfere in the communication process within the group was the use of testimonials. A prisoner would be tricked into writing out material during interrogation. This material would then be shown to another prisoner with the allegation that it was spontaneously given.

One key technique to social alienation was to put a prisoner in a cell with four or more prisoners who were more advanced on "thought reform." The progress of the cell was defined as the progress of the recalcitrant member. The members of the cells were to keep after him until he "saw the light."

The author gives other examples from the Communist prisons. He then examines other authoritarian institutions such as prisons, mental hospitals, basic training centers, etc., and shows how they contribute to the destruction of the individual's self image. (*AC*)

AVERBACH, E. AND CORIELL, A. S. **Short-term memory in vision.** *Bell Syst. Tech. J.*, 1961, 30, 309–328.

This article is a very clear and well-written report of a series of experiments designed to investigate the functional properties of storage and erasure of information in the visual system. A random selection of 16 letters is presented to the subject in a 2 x 8 array for 50 ms. Then after a variable delay, a bar marker is presented for 50 ms above one of the letter positions of the top row or below the bottom row. Each trial consists of a newly chosen array and a randomly selected bar location. The subject is required to report the name of the letter marked. As the delay increases, the per cent correct responses decreases approaching an asymptote between 25% and 35% correct when the delay is about 150 ms. The decrease is attributed to memory decay and the residual to a transference to more permanent memory.

In the second experiment, a circle around the letter position is used instead of the bar above or below it. At negative, zero, or

long positive delays, the performance is very similar to that with the bar marker but for delays of about 100 ms a notable decrease in accuracy is observed using the circle. This decrease can be attributed only to "interference" or more properly to "erasure" of the previous signal.

The third experiment was designed to measure the time required to "read out" the marker location and letter designated. The array and bar marker were presented simultaneously and after a delay the "erasing" circle. If the erasing circle is delayed about 250 ms or more it has no effect. As the delay is decreased below this, there is a marked decrement in ability to name the letter. By simple, but sophisticated operations, it is possible to derive the "storage time" of the visual system—about 300 ms and estimate the "storage capacity" —40–50 bits. The results and conclusions of this experiment should have far-reaching effects, particularly in forced pace disjunctive reaction tasks. (*CCF*)

HOFFMAN, P. J. **The paramorphic representation of clinical judgment.** *Psychol. Bull.*, 1960, 57, 116–131.

The author outlines methodological refinements for the mathematical description of (static) judgments. The situation in which his models may be employed is one in which a judge, given a set of cues (e.g., a Rorschach profile), makes certain predictions to some criterion (e.g., psycho-neurotic trend). The author indicates how, in the past, endeavors to describe such judgments have been to use β coefficients of the multiple regression formula as indicators of the emphasis attached to each of the predictor variables by the judge. The author then points out four shortcomings of this procedure: (1) that different judgments with different multiple R's cannot be meaningfully compared in terms of the β weights; (2) that β weights do not account for all the predictable variance; (3) that β weights do not allow for the assessment of the independent contribution of each predictor; and, most importantly, (4) that the multiple regression model assumes a linear rather than a configurational form of judgment on the part of the judge. The term "configurational" implies that cues are not independent dimensions for the judge, but that, say, cue A has meaning only if accompanied by cue B. The author takes care of the first three problems by proposing a new coefficient, called the relative weight, which is based on independent components of variance. The author then discusses various ways of analyzing configurational forms of judgments and provides examples of his work on linear as well as configurational judgments. (*JWG*)

ROSENBERG, M. J., HOVLAND, C. I., MCGUIRE, W. J., ABELSON, R. P., AND BREHM, J. W. **Attitude organization and change: An analysis of consistency among attitude components.** New Haven, Conn.: Yale Univ. Press., 1960. 239 pp.

Attitudes are analyzed into cognitive, affective, and behavioral domains (Ch. 1) and models of equilibrium or consistency within and between domains are developed and tested. Particular attention is devoted to the effects of changing one attitude structure on changes in other structures logically or affectively related to it. A wide range of change techniques is employed, including hypnosis (Ch. 2), Socratic questioning (Ch. 3), and various discrepant information inputs (Ch. 4, 5). Over-all results (Ch. 6) follow:

1a. When the affective component of an attitude is altered, cognition undergoes corresponding changes. (1b) These changes persist as long as changed affect persists, but, (1c) some cognitive effects persist after affect-changes forces (e.g., post hypnotic affective state) are withdrawn.

2. Evoking an individual's own attitude inconsistencies by Socratic questioning, without supplying any new information, results in greater consistency.

3a. Persons can be persuaded more easily by messages which increase consistency, than by messages which decrease it, but, (3b) changes on related or derivative issues produced by such messages are less than expected.

4. The preferred solution to a belief

dilemma is one involving the least effortful path, one which involves the fewest sign changes in an attitude structure.

5. In resolving cognitive discrepancies, people seek not only the attainment of balance and consistency but also the solution that maximizes potential gain and minimizes potential loss.

6. Commitment to a relatively great amount of discrepant behavior will increase resistance to belief of a nonsupporting communication and will increase acceptance of a supporting communication.

7. A disliked event outside the individual's control will result in positive change toward that event when the individual feels he could have avoided it or when he feels he did have a prior choice.

It is also suggested that unresolvable disequilibria tend to become unconscious. (*WPL*)

BROWNSON, HELEN L. **Research on handling scientific information.** *Science*, 1960, 132, 1922–1931.

Out of the scientist's need to cope with a rapidly increasing amount of published scientific material, a new field of interdisciplinary research has grown in the last 10 to 15 years. This field, for which a satisfactory name has not yet been found, includes research on the storage, dissemination, and processing of scientific and technical information. Its main goals are the improvement of communication among scientists and a better utilization of accumulated scientific knowledge. Abstracting, indexing and classification of scientific literature, mechanization of information retrieval, machine translation, development of character-recognizing devices, are the main objects of research in the field. The present article is a review of the most recent developments in these areas.

An understanding of the inadequacies in the present flow of scientific information presupposes a thorough knowledge of the ways in which scientists communicate and their information needs. From the present review it appears that, although many studies have been concerned with particular aspects of the problem, a broad analysis of the scientific communication system has not yet been attempted.

In the area of information storage and retrieval, several partially mechanized systems have been developed and already are in operation in various government agencies, industrial organizations, and universities. Numerous research groups are concerned with improving these systems and experimenting with new techniques. The new computer technology offers ample opportunities for developing revolutionary methods to handle scientific information, and several research programs are now devoted to the exploitation of such opportunities. Closely interrelated with these efforts are attempts to find new procedures for indexing and classifying the subject content of scientific documents, in the hope of discovering basic patterns which could be used to handle, in a mechanized system, the semantic content of the literature. In spite of intensive efforts in this direction, research on the application, on a large scale, of machines to the handling of scientific information is still in its infancy. The prevalent feeling is that, in order to make progress in this area, the essential problems of analyzing the subject content, meaning, and relevance of documents first have to be solved.

Up to a few years ago, the main efforts of the researchers aimed at discovering ways of mechanizing only the more routine aspects of information handling, but were still relying on human judgment for the analysis of the content of written records. More and more researchers now believe that, for a large-scale attack on the problem of dissemination, storage, and retrieval of information, the possibilities of using machines to analyze *directly* the language of documents should be explored. Oriented toward this goal are research programs concerned with the automatic analysis of texts, such as the studies on indexing and abstracting techniques based on word-frequency statistics (auto-abstracts), or on the mechanization of syntactic analysis of Engish sentences followed by a reduction to simpler constructions ("transformational analysis"), or

on mechanical translation of texts from one language into another. Evaluations of many of these procedures are in progress to determine their effectiveness and efficiency in information handling. (*RT*)

LUSE, S. A. AND HARRIS, B. **Brain ultrastructure in hydration and dehydration.** *Arch. Neurol.*, 1961, 4, 139–152.

With the help of electron microscopy, the authors have studied experimental hydration and dehydration in the brain of adult albino rabbits. The brain was shrunken by intravenous hypertonic solutions. In five rabbits the cortex was biopsied after dehydration and then after the biopsy the cerebrum was hydrated with intravenous distilled water.

In the cases of edema, irrespective of the method used to produce the edema, the electromicroscopical picture was the same. The brain was swollen and there was a tremendous increase of oligodendroglial cytoplasma. In dehydration, the opposite picture was seen, the brain was shrunken and the cytoplasma of the oligodendroglia was greatly reduced. Brain hydration and dehydration seems to be primarily a function of the oligodendroglial cytoplasma and not a function of the extra-cellular space.

The article is illustrated with beautiful electron microscopic photographs of the experiments. (*WHvH*)

One might even, I believe, go so far as to say, without being paradoxical, that every biologist, however faithful to mechanistic theories, is an unconscious vitalist. Why, after all, should he have chosen this strange profession, which consists in scrutinizing the evidences of vitality, if he did not attribute, at least deep down in his subconscious mind, a privileged value to the living thing? It is in proportion as he attaches value to the vital that he strives obstinately to reduce it, and his ambition to possess it is the measure of his respect for it.

JEAN ROSTAND, *Can Man Be Modified?*

ABOUT THE AUTHORS

Nancy S. Anderson ("Comments on the Use of Computers in Psychological Research") received her B.A. in mathematics from the University of Colorado and her M.A. and Ph.D. degrees in psychology from Ohio State University. Now an assistant professor of psychology at the University of Maryland, Dr. Anderson tells of her previous work experiences and interest in computers: "In addition to a mathematics background, my work at the U. S. Navy Electronics Laboratory gave me an introduction to the use of computers. While at IBM for three years, I became experienced in the uses and applications of computers in the physical as well as the behavioral sciences." She is the author of several publications which reflect her work and interest in quantitative methods in psychology, human learning, visual perception, and engineering psychology.

John E. Coulson ("Introduction: Recent Areas of Application"), who served as chairman of the computer symposium printed in this issue, tells us that "the original suggestion that there should be a symposium on the use of computers in psychological research was made by Dr. Theodore Vallance, head of the Division which sponsored this meeting at the 1960 American Psychological Association meetings." Dr. Coulson, human factors scientist, research directorate, System Development Corporation, attended the University of Arizona where he received a B.A. in psychology and a B.S. in physics, and Columbia University where he earned his M.A. and Ph.D. degrees in psychology. At the System Development Corporation, he is serving as head of the Automated Teaching Project and is involved in research on automated instruction and investigation of educational data processing systems. He is the author of a number of articles on problems and experiments concerned with automated teaching.

William K. Estes ("A Descriptive Approach to the Dynamics of Choice Behavior"), research professor of psychology at Indiana University, writes: "I first became actively interested in problems of choice behavior while participating in a summer seminar on Decision Processes held in Santa Monica in 1952 and co-sponsored by The University of Michigan and RAND Corporation. Prior to that, my theoretical work had been exclusively concerned with simpler forms of conditioning. Since then, continuing interaction with Jacob Marschak, Clyde Coombs, and, more recently, with Patrick Suppes and Gordon Bower, has been important in stimulating my continued interest in the formulation and testing of models for choice behavior." Dr. Estes's research on learning theory and mathematical models in psychology has resulted in a number of publications, among them a book (with R. R. Bush) on *Studies in Mathematical Learning Theory*. He has had a series of articles in the *Psychological Review* and *Psychometrika*, running from 1950 to 1961, devoted to developing a statistical approach to learning theory, and has published many experimental articles concerned with punishment, human probability learning, and all-or-none phenomena in associative learning. Dr. Estes is a graduate of the University of Minnesota, where he also received his Ph.D., and is currently serving as an associate editor of the *Journal of Experimental Psychology*.

Wilson H. Guertin ("Medical and Statistical-Psychological Models for Research in Schizophrenia") is presently associate research professor in the College of Nursing and associate professor in the department of psychology at the University of Florida. An undergraduate of Rutgers University, he went on to the University of Iowa, receiving an M.A. in psychology there, and then to Michigan State University where he was awarded his Ph.D. in psychology. The author of many publications dealing with psychological tests and factor analytic evaluation of schizophrenic symptoms and syndromes, Dr. Guertin has a methodological interest in deriving concepts in the behavioral sciences from actual data, usually by employing factor analysis in large scale data processing.

A. Paul Hare ("Computer Simulation of Interaction in Small Groups") is associate professor in the department of sociology and anthropology at Haverford College. The author of the forthcoming *Handbook of Small Group Research*, he reports that he is "currently doing research on the relationship between personality variables and seating preference in small group discussions. I am also continuing to write programs for the IBM 650 to simulate interaction in small groups." Dr. Hare, who has published before in *Behavioral Science*, received a B.A. from Swarthmore and a B.S. from Iowa State College, an M.A. from the University of Pennsylvania, and a Ph.D. from the University of Chicago.

Howard Maclay ("Cross-Cultural Use of the Semantic Differential") received his B.A. from the University of Chicago and his M.A. and Ph.D. degrees from the University of New Mexico. Now research associate professor at the Institute of Communications Research of the University of Illinois, Dr. Maclay writes: "After taking a Ph.D. in anthropology with a concentration in linguistics at the University of New Mexico, I studied experimental psychology and quantitative methods at Harvard under a Social Science Research Council postdoctoral fellowship (1955–56) and served as a research fellow on the Southwest Project in Comparative Psycholinguistics which studied language behavior among a number of American Indian groups. I also participated in the interdisciplinary conference on the behavioral sciences held by the Air Force Office of Scientific Research in New Mexico during the summers of 1957–58." He is presently engaged in research on interpersonal communication involving linguistic messages, and on the application of experimental and quantitative techniques to problems in cultural anthropology and linguistics. The author of many publications in these interest areas, his most recent work is entitled "A Descriptive Approach to Communication" which will appear in a book to be edited by S. Washburne and published by Pergamon Press.

William N. McPhee (Abstract of "Empirical Inputs in a Voting Model") is research associate in the Bureau of Applied Social Research and lecturer in the graduate department of sociology at Columbia University. Dr. McPhee, who attended Yale as an undergraduate and received his Ph.D. from Columbia, is interested in research on models of mass behavior.

Richard L. Meier ("Explorations in the Realm of Organization Theory IV: The Simulation of Social Organization") reports that "This review was assembled at the time we undertook to develop a series of subcomputer level simulations emphasizing various facets of organization." The recipient of a Ph.D. from U.C.L.A., Dr. Meier is research social scientist at the Mental Health Research Institute and associate professor of conservation in the School of Natural Resources at The University of Michigan. His book *Science and Economic Development* reflects his work on studies of social, technological, political, and other features of economic development. Other research is underway on problems of social organization, and in particular he is concerned with the measurement of flow and accumulation of social communications. The author of more than 40 articles and reports, his most recent book is *Modern Science and the Human Fertility Problem.*

Lucian W. Pye ("Personal Identity and Political Ideology"), professor of political science and senior staff member of the Center of International Studies at M.I.T., received his Ph.D. from Yale University after undergraduate work at Carleton College. His interest in personality change and political development in newly emergent countries is evident in his study, now nearing completion and tentatively entitled *Politics, Personality and Nation Building*, of some of the psychological inhibitions to economic development and nation building, based upon field research in Burma. Other books which reflect Dr. Pye's work in this area are *Guerrilla Communism in Malaya*, a study of the appeals of Communism in an underdeveloped area, based on depth interviews with former communists; *The Politics of the Developing Areas*, and the forthcoming *The Emergent Nations*. The author of numerous articles which have appeared primarily in political science journals, he has been an active member of the Comparative Politics Committee of the Social Science Research Council and is now organizing a conference on "Communications and Political Development" for that committee. In the field of public affairs, he has lectured to the war colleges and has been identified as a member of President Kennedy's original "brain trust." With regard to his research, he writes: "Recently, I have become increasingly interested in the various ambivalences which seem to underlie the attitudes of people in transitional societies toward modern life, and particularly toward economic development and national strength. The popular American versions of a 'revolution of rising expectations' grossly oversimplifies the extremely complex fears and anxieties such people can have about their individual and collective identities."

Richard N. Rosecrance ("Categories, Concepts, and Reasoning in the Study of International Relations") writes that this article "stems from several years of ruminating about the methodological apparatus of international relations. I was first stimulated to think along these lines by A. N. Prior's *Logic and the Basis of Ethics*. It occurred to me that certain of the criticisms he made of the proponents of the 'naturalistic fallacy' in ethics should be made of practitioners of international study. More important, perhaps, I became convinced that certain uses of terms, even if leading to *synthetic* propositions, were dubious on other grounds. Even *synthetic* definitions could effectively

prejudice the empirical issue in advance. These thoughts led to a correspondence with Professor Karl Deutsch lasting a year or more; the final form of this article benefits very much from that exchange and owes much to his criticisms. I should add that I have recently completed the manuscript of a book entitled *Empirical Systems of International Relations* which seeks to remedy certain of the difficulties mentioned in the article." Educated at Swarthmore College (B.A.) and Harvard University (Ph.D.), Dr. Rosecrance is assistant professor of political science and co-ordinator of the national security studies program at the University of California, Los Angeles. Regarding his current work, he goes on to say "My present research is in the special field of national security policy and in the general field of theory of international relations. I am also engaged in a research project under the general editorship of Professor Louis Hartz of Harvard on 'The Foundations of Australian Society'". Dr. Rosecrance is the author of a recent book *Australian Diplomacy and Japan: 1945–1951*.

Harry F. Silberman ("A Computer-Controlled Teaching Machine"), human factors scientist at the System Development Corporation, is a previous contributor to *Behavioral Science*. He took his undergraduate training in biology at Chico State College and received his M.A. and Ed.D. degrees from U.C.L.A. in educational psychology. With Dr. Coulson, he is a member of the Automated Teaching Project being carried on at System Development Corporation, where he is interested in the application of computers in educational research and in principles underlying auto-instructional methods. Dr. Silberman is a co-author with Dr. Coulson of several articles.

William H. Starbuck ("Testing Case-Descriptive Models"), an instructor in the School of Industrial Management at Purdue University, relates that "this paper grew out of discussions with Geoffrey Clarkson, Kalman Cohen, and Julian Feldman, and is concerned with problems of the type which they encountered in their research. Herbert Simon has proposed that a case-descriptive model be tested by giving a group of knowledgeable judges both the raw data and the output of the model, and asking them to decide which is data and which is output. If the judges cannot distinguish between the data and the output, the model would be sufficient. The testing procedure suggested here is similar to Simon's proposal except that (1) the model itself defines the characteristics of a judge, and (2) the model is sufficient if even one judge mistakes output for data." Mr. Starbuck is a Ph.D. candidate at the School of Industrial Administration, Carnegie Institute of Technology, where he received his M.S. degree in industrial administration after earning his A.B. in physics at Harvard. He is currently interested in research in mathematical models of social behavior.

Edward E. Ware ("Cross-Cultural Use of the Semantic Differential"), research assistant professor, Institute of Communications Research, University of Illinois, reports that he is "presently engaged in a research project concerned with the development and validation of a form of semantic differential for measurement in the personality area. I am particularly interested in the investigation of individual differences in the frames of reference involved in interperson perception and in the application of factor analysis and multivariate procedures in this type of investigation." A native of Richmond, Virginia, Dr. Ware received his undergraduate education (and his M.A. degree) at the University of Richmond, after which he went to the University of Illinois where he gained his Ph.D. in psychology.

Benjamin W. White ("The Computer as a Pattern Generator for Perceptual Research") is a staff member of the psychology group at Lincoln Laboratory, Massachusetts Institute of Technology. Of his work at Lincoln Lab, he writes: "An interest in perception stemming from undergraduate work with Hans Wallach at Swarthmore has found fertile soil at Lincoln where there are people from many disciplines working on problems of pattern recognition, and where computers not only live, but are born. The computer has proved to be a useful tool with which to generate complex stimuli for the study of pattern recognition and the kinetic depth effect, not to mention the promise it shows as a simulator of the pattern recognition process." Dr. White, who received his A.B. from Swarthmore College and his Ph.D. from The University of Michigan, is engaged in research on human pattern recognition, the use of computers to generate stimulus patterns, auditory and tactual outputs for reading machines, and visual and auditory perception. He has a number of recently published articles in these research areas.

THE INTERNATIONAL JOURNAL OF SOCIAL PSYCHIATRY

Vol. VI, Nos. 3 & 4 **Autumn 1960**

This interdisciplinary journal, covering all aspects of the various social sciences, is of interest to psychiatrists, psychologists, sociologists, educationists, criminologists, mental health officers, anthropologists, and other workers in the field of mental health and social services. The second part of the special anniversary volume has now been published and contains articles by the following authors:

GENERAL:

Joshua Bierer; Robert H. Felix; D. Müller-Hegemann.

PROBLEMS OF DELINQUENCY:

A. C. Pacheco e Silva; D. H. Stott; Eleanor T. Glueck.

CULTURAL PROBLEMS:

Theodora M. Abel; Rhoda Métraux; Ilza Veith; Eduardo D. Maldonado Sierra, Richard D. Trent and Ramon Fernandez Marina; C. H. Hardin Branch and Robert Anderson; José Angel Bustamante.

RESEARCH:

Elaine R. Gladston and Ginette Amado-Haguenauer; Glenn T. Strand and William R. Larson; Asya L. Kadis and Charles Winick; Lynn Gillis; A. Mateo Alonso.

Reports Book Reviews

SUBSCRIPTION RATE (Four Issues)

£2.0.0. or $8.00 for Hospitals, Libraries, Universities, etc.
£1.5.0. or $5.00 for Individual Subscribers.

All enquiries should be addressed to

The Secretary,

The International Journal of Social Psychiatry,
7 Hollycroft Avenue, London, N.W. 3.